Spring 1-1-2013

Forward and Inverse Modeling of GPS Multipath for Snow Monitoring

Felipe Germia Nievinski

Forward and Inverse Modeling of GPS Multipath for Snow Monitoring

by

Felipe Geremia Nievinski

B.E., Universidade Federal do Rio Grande do Sul, 2005

M.Sc.E., University of New Brunswick, 2009

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Aerospace Engineering Sciences

2013

This thesis titled:

Forward and inverse modeling of GPS multipath for snow monitoring

written by Felipe Geremia Nievinski

has been approved for the Department of Aerospace Engineering Sciences

_____

Kristine M. Larson

_____

Dennis Akos

_____

George Born

_____

Eric Tilton

_____

Mark Williams

_____

Valery Zavorotny

Date: _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Nievinski, Felipe G. (Ph.D., Aerospace Engineering Sciences)

Forward and inverse modeling of GPS multipath for snow monitoring

Thesis directed by Professor Kristine M. Larson

Snowpacks provide reservoirs of freshwater, storing solid precipitation and delaying runoff to be released later in the spring and summer when it is most needed. The goal of this dissertation is to develop the technique of GPS multipath reflectometry (GPS-MR) for ground-based measurement of snow depth. The phenomenon of multipath in GPS constitutes the reception of reflected signals in conjunction with the direct signal from a satellite. As these coherent direct and reflected signals go in and out of phase, signal-to-noise ratio (SNR) exhibits peaks and troughs that can be related to land surface characteristics. In contrast to other GPS reflectometry modes, in GPS-MR the poorly separated composite signal is collected utilizing a single antenna and correlated against a single replica. SNR observations derived from the newer L2-frequency civilian GPS signal (L2C) are used, as recorded by commercial off-the-shelf receivers and geodetic-quality antennas in existing GPS sites. I developed a forward/inverse approach for modeling GPS multipath present in SNR observations. The model here is unique in that it capitalizes on known information about the antenna response and the physics of surface scattering to aid in retrieving the unknown snow conditions in the antenna surroundings. This physically-based forward model is utilized to simulate the surface and antenna coupling. The statistically-rigorous inverse model is considered in two parts. Part I (theory) explains how the snow characteristics are parameterized; the observation/parameter sensitivity; inversion errors; and parameter uncertainty, which serves to indicate the sensing footprint where the reflection originates. Part II (practice) applies the multipath model to SNR observations and validates the

resulting GPS retrievals against independent *in situ* measurements during a 1-3 year period in three different environments – grasslands, alpine, and forested. The assessment yields a correlation of 0.98 and an RMS error of 6-8 cm, with the GPS under-estimating *in situ* snow depth by approximately 15%. GPS daily site averages were found effective in mitigating random noise without unduly smoothing the sharp transitions as captured in new snow events. This work corroborates the readiness of quality-controlled GPS-MR for snow depth monitoring, reinforcing its maturity for operational usage.

Acknowledgments

"It takes a village to raise a child," says the African proverb. Likewise, it takes many people to form a PhD graduate. I am indebted first and foremost to my academic advisor, Dr. Kristine M. Larson, a real trailblazer for whom I hold the deepest admiration. Your scientific excellence and integrity are truly exemplary. I will remember your relentless pragmatism and exacting standards with appreciation.

I could not have produced a work of this depth and breadth from the ground up. Larson's vision in pioneering the development of GPS multipath reflectometry (GPS-MR) can be equated to uncovering the tip of a rich vein of ore, that could be mined on and off for many years. Notwithstanding the above initiation – sometimes it felt as if all I was doing was "dotting the i's and crossing the t's" – any faults that remain are obstinately my own. I am confident that the theoretical and experimental foundations established priorly and presently offer a fertile ground upon which many fruitful developments are sure to follow. I feel fortunate to have become part of this research community around GPS-MR – clearly an idea whose time has come.

I would like to extend my gratitude to the members of my dissertation committee, for being generous in sharing their time and expertise with me. Dr. Dennis Akos, for demystifying how a GPS receiver can be made from scratch and for letting me play with an 18-m parabolic antenna. Dr. George Born, for the Kalman filter lessons and the space race anecdotes. Dr. Eric Tilton, for exposing me to the larger hydrological applications of GPS-MR and for the perspicacious dissertation comments. Dr. Mark Williams, for teaching me all I know about snow and helping an unlikely Brazilian to be at ease in a snowpit. Last but certainly not least, Dr. Valery Zavorotny, for clarifying many electromagnetic scattering intricacies; your unselfish contribution to my incipient knowledge will not be forgotten.

Next I wish to give thanks to Dr. Glenda Russell, for facilitating the Dissertation Support Group, as well as to fellow participating graduate students. This group infused a sense of camaraderie in the hardship of finishing a PhD dissertation and bolstered my self-confidence to become a fluent academic writer.

Numerous current and past colleagues and collaborators in the GPS Reflections group have offered invaluable feedback and cooperation during the years: Dr. John Braun, Dr. Karen Boniface, Dr. Staffan Backen, Dr. Ethan Gutmann, Dr. Andria Bilich, Dr. James McCreight, Dr. Evan Pugh; Qiang Chen, Clara Chew, John Pratt, Sarah Evans, Maria Rocco, Wei Wan, Kyle Wolma, Natasha Whitney, Chen Zhao, and Praveen Vikram.

Outside the Boulder area, I acknowledge Dr. Sergio Florencio de Souza (UFRGS) for serving as a tutor in Brazil and Sajad Tabibi (Université du Luxembourg) for testing some of the software routines developed.

I would also like to thank Dr. Penina Axelrad (UC Boulder, ASEN), Dr. Al Gasiewski (UC Boulder, EE), Dr. Fabiano Rodrigues (UT Dallas), Dr. Marcelo Santos (UNB), Dr. Yara Regina Marangoni (USP, IAG), Dr. Edvaldo Simões da Fonseca Junior (USP, EP), Dr. João Carlos Chaves and Dr. João Francisco Galera Monico (UNESP). They have offered me opportunities to present versions of this work in progress and have provided encouragement to further advance it.

On a personal level, my whole family has always been there when I needed. Thank you mother, Maria Geremia Nievinski, for all the soothing advice; father, Nelson Pedro Nievinski, for the closeness and the jokes; sister, Paloma Geremia Nievinski, for putting up with my own jokes; and brother, Henrique Geremia Nievinski, for the conversations. Giovana Ferrer Vazatta, my wife whom I adore: you have been the ideal companionship – thank you for all the shared

moments of joy and for putting up with the mess that a two graduate student household can become sometimes. *Obrigado a todos.*

Portions of this work used data provided by the Niwot Ridge Long-term Ecological Research Site, the International GNSS Service, and UNAVCO, the latter of which is also cordially acknowledged for permitting my last-minute participation in training courses on GPS and InSAR. Funding has been provided by NASA, the Fulbright Program, and the Brazilian Capes; the AGU and UNAVCO have provided additional travel support.

Table of Contents

List of Tables

List of Figures

List of Symbols

SNR; $\mathrm{tSNR}, \mathrm{dSNR}$	signal-to-noise ratio; its trend and detrended fringes
P; P_d, P_r, P_c, P_i, P_e, P_n; P_r^C, P_r^I	power; direct, reflected, composite, interferometric, error, noise; coherent, incoherent reflected power
V; V_d, V_r, V_c, V_i, V_e	voltage; direct, reflected, composite, interferometric, error
ϕ; ϕ_d, ϕ_r, ϕ_c, ϕ_i, ϕ_e	voltage phase; direct, reflected, composite, interferometric, error
$\bar{E}$; $\bar{E}_\mathrm{d}$, $\bar{E}_\mathrm{r}$; E; E^R, E^L; $E^\mathrm{L/R}$	vector electric field ($\in \mathbb{C}^2$); direct, reflected; its scalar components; RHCP, LHCP; their polarimetric ratio
ϕ^R, ϕ^L; $\phi^\mathrm{L-R}$	RHCP, LHCP electric field phase; their polarimetric difference
Φ; Φ^R, Φ^L; $\Phi^\mathrm{L-R}$; Φ_d, Φ_r ; $\vec{\Phi}$; $\widetilde{\Phi}$	antenna phase; RHCP, LHCP components; their polarimetric difference; evaluated at the direct, reflected directions; phase center offset (spatial vector $\in \mathbb{R}^3$); phase center variation (scalar)
G; G^R, G^L; G_d , G_r	antenna gain; its RHCP, LHCP components; evaluated at the direct, reflected directions
$\bar{L}$, L, L_iso; L^R, L^L; $\bar{L}_\mathrm{d}$, $\bar{L}_\mathrm{r}$	antenna complex vector effective length ($\in \mathbb{C}^2$), its norm, for an isotropic antenna; RHCP, LHCP components; evaluated at the direct, reflected directions
P_d^R	signal power collected by an isotropic RHCP antenna
τ; τ_d, τ_r, τ_c, τ_i, τ_e, τ_0	propagation delay; direct, reflection, composite, interferometric, error, tracked
Y; Y_d, Y_r	pre-correlation voltage; direct, reflected
W, Λ, Γ, Ψ; $\delta\tau, \delta f$; W_d, W_r, W_c	ambiguity function, code correlation function, normalized sinc function, phase-lock complex factor; delay, frequency tracking differences; evaluated for the direct, reflected, composite signals
T; $T_\mathrm{A}, T_\mathrm{R}$	noise-equivalent temperature; antenna, receiver
N_0, B_n	noise power spectral density; noise bandwidth
$\gamma, \lvert\gamma\rvert$	complex coherence; degree of coherence
S^2; s_h	random surface roughness effect on coherent power; surface height standard deviation (with respect to trend surface)
X, ϕ_X; X^R, X^L	coupled surface/antenna coefficient, its contribution to the interferometric phase; RHCP-, LHCP-producing coefficients.
R; R^X, R^S; R^H, R^V; $\bar{\bar{R}}$; $\bar{\bar{P}}^\mathrm{X}$, $\bar{\bar{P}}^\mathrm{S}$; ϵ	Fresnel reflection coefficients; cross- and same-sense circularly polarized; horizontal, vertical linearly polarized; reflection matrix ($\in \mathbb{C}^{2\times2}$); cross- and same-sense polarizing matrices; permittivity

$K, K_{\mathrm{dB}}, K_{\mathrm{dB}}^{(j)}$	noise bias: power effect, in decibels, its polynomial coefficients
$B\colon \lvert B\rvert^2, B_{\mathrm{dB}}, B_{\mathrm{dB}}^{(j)}$; $\phi_B, \phi_B^{(j)}, \varphi_B, \hat{\phi_B}$	reflection bias: complex-valued, power factor, in decibels, its polynomial coefficients; full phase bias, its polynomial coefficients, constant term (phase-shift), higher-order remainder
$H; H_A, H_B, H_S, H_T$	reflector height; *a priori* value, unknown bias, snow-related, topographical
e, a; $\hat{\rho}$; $e_{\mathrm{d}}, a_{\mathrm{d}}, \hat{\rho}_{\mathrm{d}}$; $e_{\mathrm{r}}, a_{\mathrm{r}}, \hat{\rho}_{\mathrm{r}}$	elevation angle, azimuth; direction (unity spatial vector $\in \mathbb{R}^3$); direct, reflected
λ, f, k	carrier wavelength, frequency (in hertz), wavenumber
X, Y; x, y; $\mathrm{C}, \mathrm{C_x}, \mathrm{C_y}$; Σ, σ, ρ; J	linear algebra vector of parameters and of observations; parameter corrections and observation residuals; covariance matrix, of parameters and of observations; vector of standard deviations, standard deviation, correlation; Jacobian matrix
$k_z; Q; \sigma_{\mathrm{e}}$	vertical wavenumber; height ramp or quadratic phase; mean extinction

Chapter 1: Introduction

Snowpacks provide reservoirs of freshwater, storing solid precipitation and delaying runoff, to be released later in the spring and summer when it is most needed (Armstrong and Brun 2008). One sixth of the world depends critically on this resource for drinking, irrigation, power generation, transportation, and recreation (Barnett et al. 2005). On the other hand, snowmelt has caused disastrous floods on many occasions (Eamer 2007). For weather and climate studies, knowledge of the snowpack is important to close the surface energy balance budget, through exchanges with the overlying atmosphere (Lemke et al. 2007). Thus both scientists and water supply managers need to know how much freshwater is stored in the snowpack and how fast it is being released as a result of melting (Rott et al. 2010).

The need for snow measurements is acknowledged in the National Research Council's Decadal Survey (NRC 2007). For example, among its consensus recommendations for the future is the Snow and Cold Land Processes (SCLP) spaceborne mission, aimed at mapping snow accumulation for fresh water availability. Greatly complementary to such spaceborne sensors are ground-based methods; the latter not only serve as essential independent validation and calibration for the former, but are also valuable for climate studies and flood/drought monitoring on their own. The goal with this contribution is to consolidate the technique of GPS multipath reflectometry (GPS-MR) for ground-based measurement of snow.

GPS multipath stands for the simultaneous reception of multiple radio signals that have propagated through different paths after their transmission by a GPS satellite (Figure 2). The least-time or shortest-distance signal is said to travel on the direct path, normally associated with the line-of-sight, close to the straight-line joining satellite and receiver. The remaining signals arrive delayed after being scattered off an intervening medium. Scattering can originate in the

atmosphere (neutral or ionized particles and layers), electronic components (loading mismatch), antenna installation (satellite body or ground monument), the natural Earth (land, water, vegetation, etc.), or the built environment. The receiving platform can be deployed at various altitudes, from near-surface (few meters tall) to elevated (towers, cliffs), all the way up to low-Earth orbit. Here we focus on GPS multipath involving reflections off large natural surfaces and recorded with near-surface receivers.

Although multipath is detrimental for GPS positioning applications, its benefits for reflectometry have become increasingly clear in the last decade. More specifically, the frequencies and amplitudes of the multipath modulation observed in GPS data show strong correlations with environmental characteristics such as soil moisture, snow depth, and vegetation growth (Larson et al. 2008; Larson et al. 2009; Rodriguez-Alvarez et al. 2011; Small et al. 2010). So while a geodesist uses a multipath-suppressing antenna installed upright, someone more interested in remote sensing would prefer to tip the antenna such that its boresight faces the Earth.

Here we further concentrate on data collected at GPS base stations, using commercial off-the-shelf stand-alone receivers and geodetic-quality antennas. Although this type of instrumentation is not ideal for reflectometry purposes, it has been demonstrated feasible and adequate in the past (Larson et al. 2009). The advantage of such a selection is that it enables us to leverage the existing continuously-operating GPS infrastructure (see Figure 1) – hundreds of ground-based receivers and 30+ L-band spaceborne transmitters – deployed and maintained at no additional cost. As the future of the GPS constellation – and similar GNSS from other nations – is secure, there are good prospects for 120 satellites in 10 years. Moreover, the tendency is for the number of publicly accessible receivers to increase.

I developed a forward/inverse approach for modeling GPS multipath present in signal-to-noise ratio (SNR) observations. It is unique in that it capitalizes on known information about the antenna response and the physics of surface scattering, to aid in retrieving the snow conditions in the antenna surroundings. The physically-based forward model, founded on geometrical optics, is utilized to simulate observations. In conjunction, a statistically-rigorous inverse model based on non-linear least squares is employed to retrieve parameter corrections responsible for observation residuals. The forward model is kept as simple as possible, given the voluminous dataset available, comprising hundreds of stations tracking dozens of satellites and spanning multiple years at one-hertz sampling rates. The inverse model adopts a parameterization of unknowns that avoids under- or over-fitting.

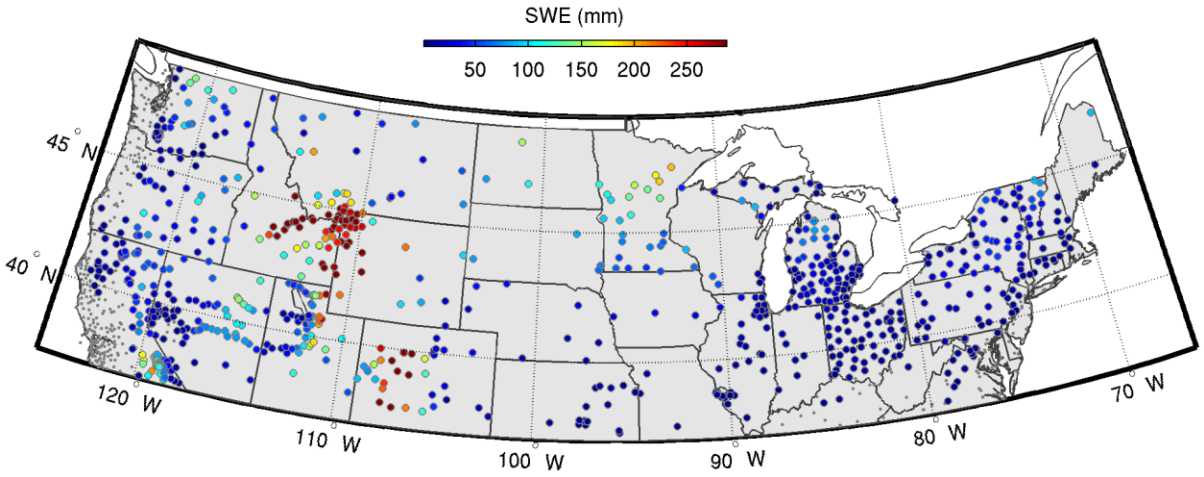

Figure 1: Continuously-operating GPS sites in the contiguous United States – both PBO, http://pboweb.unavco.org/, and CORS, http://www.ngs.noaa.gov/CORS/, networks. Color values are climatic annual snow water equivalent (SWE), (Armstrong et al. 2005); values are clipped at 97.5% for greater color discrimination; GPS sites with negligible SWE (< 1 mm) are shown in gray.

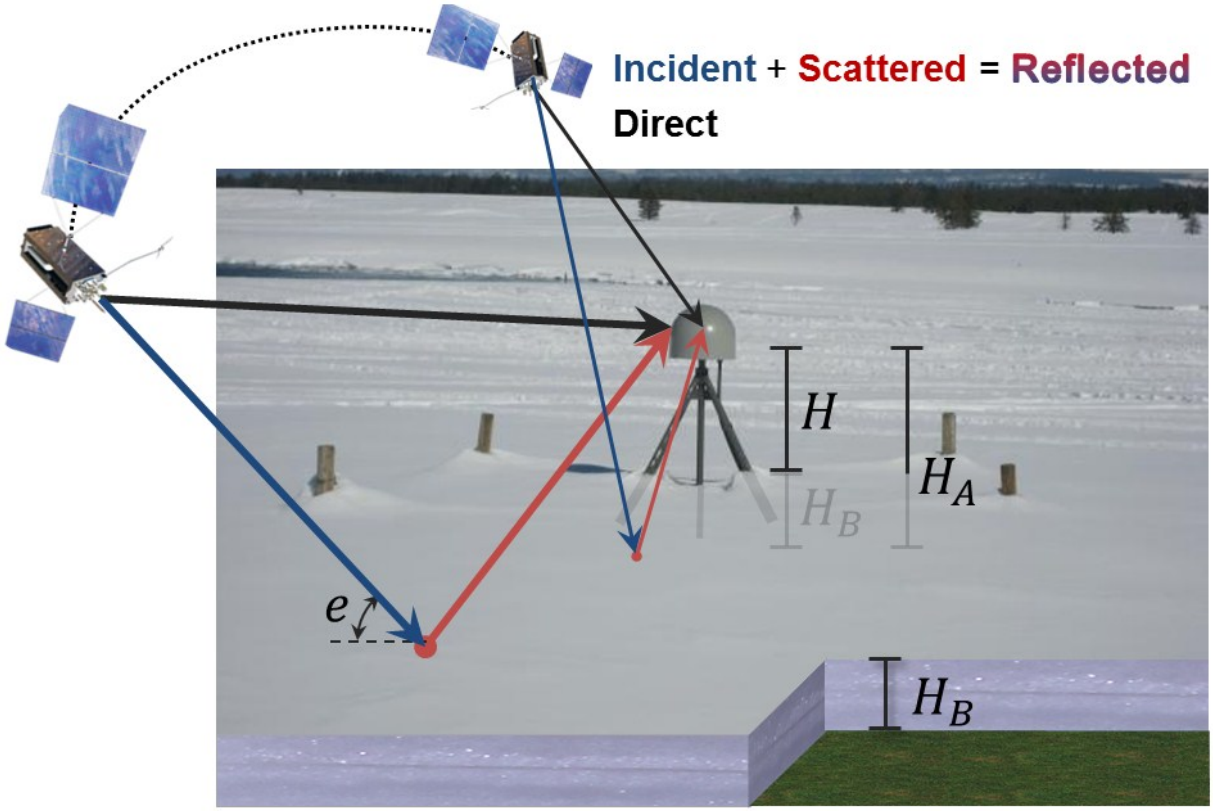

Figure 2: Standard geodetic receiver installation. The antenna is protected by a hemispherical radome. The monument (tripod structure) is ~ 2 meters above the ground. GPS satellites rise and set in ascending respectively descending sky tracks, multiple times per day. The specular reflection point migrates radially away from the receiver for decreasing satellite elevation angle. The total reflector height is made of an *a priori* value and and unknown bias driven by thickness of the snow layer.

Chapter 2: Background

2.1 SNOW MEASUREMENTS

Optical spaceborne sensors provide important information on snow extent, but cannot provide information about snow depth or snow water equivalent (SWE). These values can be measured with microwave instruments (both passive and active) (WMO 2007). The resulting coarse resolution (25 km) satellite images, although valuable in gentle terrain, are not as accurate in mountain basins, where the variable landscape incurs errors within a pixel or require significant regional tuning (WMO 2007).

In situ snow observations also come with their own limitations. Snow courses miss temporal dynamics of accumulation and ablation because the manual survey involved is carried out infrequently (typically once a month). Snow pillows and ultrasonic sensors (measuring, respectively, the weight of, and vertical distance to, the snowpack) have higher temporal resolution (typically once every 15 min) but miss important spatial variability because of their small footprint (typically 3-by-3 m for snow pillows, less for ultrasonic sensors). On a global scale, the value of *in situ* observations is often reduced by large inconsistencies in methods, frequency, and standards (WMO 2007).

Ground-based remote sensing devices, such as radars, radiometers, and scatterometers complement long-range remote sensing satellites (having a limited revisit time) and airplanes (operating sporadically), as these former sensors can function at a much higher rate and continuously in time. Furthermore, in contrast to many *in situ* probes, ground-based remote sensing sensors have a much larger spatial footprint (from ~ 50-100 m), thus being more comparable to, and valuable for, the validation of space-borne sensors, such as the AMSR-E

(Kelly et al. 2003). However, at present these sensors are not operational, being typically used in targeted campaigns such as the Cold Land Processes Experiment (CLPX) (Hardy et al. 2008).

2.2 GPS MULTIPATH REFLECTOMETRY (GPS-MR)

Among the commonly recorded GPS observables – code pseudorange, carrier phase, and signal-to-noise ratio (SNR) – only the ranging observables are routinely used for position determination. In contrast, SNR is the primary observable for multipath reflectometry. SNR exhibits peaks and troughs as direct and reflected signals go in and out of phase (Figure 3). Thus when multipath is present, it tends to be the dominating feature in SNR (besides the antenna gain pattern). Ranging observables could in principle be manipulated to extract multipath – using, e.g., carrier-phase post-fit residuals, single-frequency carrier-minus-code, dual-frequency "MP" linear combinations, etc. – but the signature revealed is typically not as clear as in SNR, and requires more laborious processing. Sometimes multiple receivers, antennas, or correlators are employed to aid in more specialized reflectometry modes, but the need for special-purpose hardware would prevent leveraging the existing GPS infrastructure.

Figure 3 shows SNR observations (measured, simulated, and their residual), for a typical ~ 2-m tall antenna, whose gain is hemispherical in elevation angle and omnidirectional in azimuth. The trend follows from the gain while the fringes result from the interference between direct and coherently reflected signals. The transmitting satellite is rising on the left and culminating near zenith on the right, after which it proceeds to set on the horizon (not shown). Among the several GPS signals broadcast, here we use the newer L2C signal, because its modulation codes are publicly known (contrary to the P(Y) signal) and are sufficiently long (in contrast to C/A), thus avoiding tracking losses that can disfigure the multipath signature in SNR.

2.3 RELATED DEVELOPMENTS

For GPS positioning applications, one seeks to suppress multipath. Several techniques have been developed with that goal. They can be roughly divided into four categories:

(1) environment: minimizing the presence of reflecting surfaces

(2) antenna: hemispherical pattern, good cross-polarization discrimination, multi-element arrays

(3) receiver: algorithms to analyze the correlation waveform (Pany 2010)

(4) observation modeling: modified sidereal filtering (Choi et al. 2004), combined code and phase observations (Satirapod and Rizos 2005), etc.

Admittedly, the complexity of most environments makes it difficult to model multipath from first principles and match the observations. This issue is compounded by the sensitivity of the model output to perturbations in the input receiver position as well as defects in the surface description. Thus forward simulations are rarely sufficiently accurate to be used as measurement corrections. Yet their error envelopes are often useful as bounds on the expected error, for performance integrity in navigation augmentation and stochastic weighting in geodesy (Aloi and van Graas 1999; Byun et al. 2002; Chen et al. 2009; Ercek et al. 2005; Evans et al. 1989; Fan and Ding 2006; Geren et al. 2008; Gomez et al. 1995; Hannah et al. 1998; Lau and Cross 2007; Lippincott et al. 1996; Macabiau et al. 1999; Mora-Castro et al. 1998; Ray and Cannon 2001; Rigden and Elliott 2006; Rodgers 1992; Weiss et al. 2007; Zhu and van Graas 2009). Also, by offering repeatable multipath conditions, a simulator serves as a neutral assessment criterion in the comparison of competing antennas and receivers (Irsigler et al. 2005; Spirent 2010) and their impact on positioning solutions (King and Watson 2010). Finally, in principle measurements could be used to estimate improved values for the underlying parameters driving the forward

simulation, leading to potentially more accurate corrections for multipath errors and positioning solutions (Bilich et al. 2008).

In parallel to efforts by the positioning community, the reflectometry community has developed ways to exploit GPS reflections for remote sensing purposes (Figure 4). Focus has been on relatively simpler environments, offering homogeneous scattering conditions, as it is questionable how feasible, reliable, and stable retrievals could be in the most complicated environs. We can identify two general classes of reflectometric methods: "coherent" and "incoherent". The fundamental difference lies in the way speckle or interference fringes are treated. It is an undesirable noise source in incoherent reflectometry, while for coherent methods it is the signal of interest. Glenson et al. (2009) provide a thorough review of incoherent reflectometry. Cardellach et al. (2012) provide a more recent review, including some coherent methods, but does not focus much on single-antenna mode, which we now address.

Typically direct and reflected signals are correlated against separate replicas. To facilitate this separation, two antennas are employed to collect the two signals. Even when only one antenna is used, an adequate delay separation is sought, by observing high elevation satellites from high altitude platforms. Alternatively, a single replica can be correlated against the poorly separated direct-plus-reflection signal. We call this mode multipath reflectometry (GPS-MR). It relies on the coherence of reflections. Of course, coherent reflections can also be exploited in dual-replica mode, e.g., differencing separately tracked direct and reflected phase. So our work is relevant for both single- and dual-replica coherent reflectometry – i.e., multipath-based (Anderson 2000; Jacobson 2010; Kavak et al. 1998; Ozeki and Heki 2011; Park et al. 2010; Rodriguez-Alvarez et al. 2011a; Treuhaft et al. 2001) or not (Belmonte Rivas and Martin-Neira

2006; Cardellach et al. 2012; Cardellach et al. 2004; Fabra et al. 2012; Helm et al. 2007; Martin-Neira et al. 2002; Rodriguez-Alvarez et al. 2011b; Semmling et al. 2011).

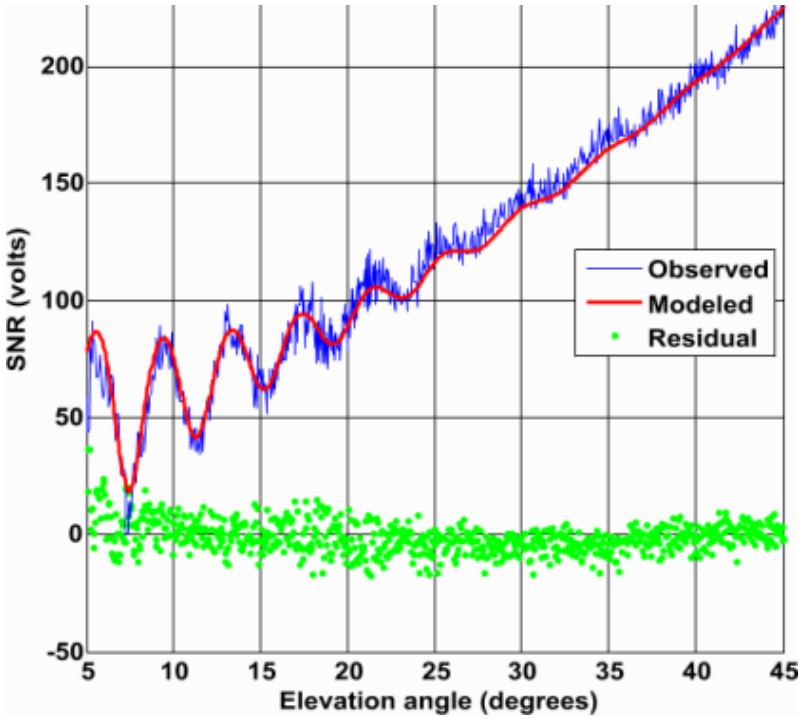

Figure 3: Inversion results (red trace) over measurements (blue) for a single GPS track, and corresponding residuals (green dots). Deep fades are consequence of the destructive interference between direct and reflected signals. Residuals, originally zero-centered, have been displaced vertically for clarity. SNR units are normalized by an arbitrary constant for clarity.

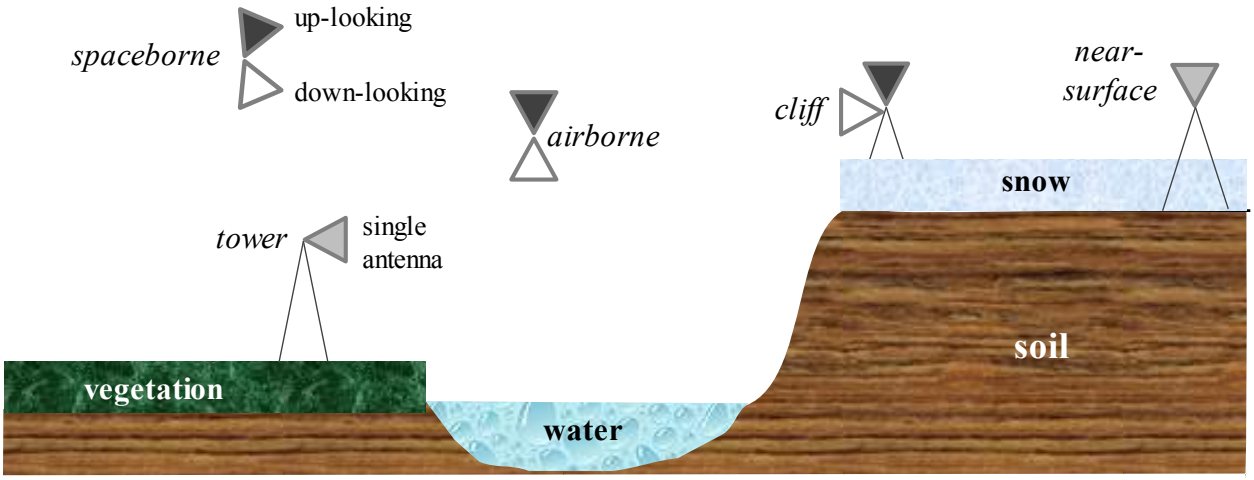

Figure 4: Various configurations and settings possible in GPS reflectometry.

Chapter 3: Forward modeling of GPS multipath for near-surface applications

3.1. INTRODUCTION

The forward model produces surface reflections as approximate replicas of the direct signal. The amount of attenuation as well as group and phase delay suffered by the reflections is dictated by the properties of target surface itself, as well as by the GPS measurement system characteristics. The former includes geometry and dielectric properties of the ground or snow media. The latter comprises antenna gain and phase patterns, signal modulation, and receiver tracking algorithms. Finally, characteristics of the monitoring setup, such as the reflector height (antenna height above the ground) and its orientation, also need to be taken into account. Given *a priori* information about each of the parameters above, the forward model furnishes simulation counterparts to the measured observations.

For reflectometry, in order to convert observed multipath parameters into useable environmental products, it is important to be able to explicitly link the GPS observables to known characteristics of the GPS receiver/antenna and the reflecting environment. A simulator can also be used for the design of new installations and feasibility studies, aiming at maximizing the reception of reflections. In positioning applications, a simulator supports multipath mitigation efforts in terms of, e.g., site selection, antenna design, receiver performance assessment, and in relating different observations to a common parameterization.

Existing GPS forward models are frequently developed for specific goals (e.g. satellite missions and/or receivers), and their respective implementations, in the form of software simulators, are generally not readily available for the general scientific community to use and/or modify. Here a fully polarimetric forward model is presented, accounting for right- and left-handed circularly polarized components of the antenna response as well as of both direct and

reflected signals. A simulator has been developed in MATLAB implementing the forward model. It can simulate L1 and L2 carrier frequencies and C/A, P(Y), and L2C modulations; it produces signal-to-noise-ratio (SNR), code pseudorange, and carrier phase observables. In addition to including specific reflecting materials (water, concrete, soil, etc.), it allows certain dimensional properties to be varied, such as soil moisture, snow density, and accounts for random surface roughness.

The forward model presented here is based on that developed by Zavorotny et al. (2010); it has been extended to allow for variable incident power and polarization, antenna orientation, antenna model, code pseudorange, and noise power. In the next section we review related contributions in the positioning as well as in the reflectometry literatures. It provides context and motivation for these forward modeling efforts. In the subsequent section we will describe each of the components of the forward model. We conclude with simulations in various reflection scenarios.

3.2 OTHER GPS MULTIPATH SIMULATORS

We identify three types of multipath simulators in the literature. Tracking simulators focus on the measurement/replica signal matching, often adopting arbitrary values for the reflection power, phase, and delay. Geometrical simulators calculate the reflection delay based on a given surface geometry, receiver position, and satellite direction; the reflection power often remains empirically defined. Polarimetric simulators account for the polarization matching between surface and antenna responses, yielding physically-based reflection power and phase values.

Each type of simulator admits subtypes. For example, tracking simulators can neglect the code modulation; implement one particular code discriminator; provide the full correlation waveform over varying delays, or even the Doppler-delay map over increasing frequency steps;

synthesize the radiofrequency signal, sampled as input to a software receiver or hardware-fed to a conventional receiver.

Geometrical simulators can support different surface shapes. Horizontal surfaces are the simplest to model; they may serve as an adequate approximation for receivers installed in open spaces. A tilted surface offers more degrees of freedom to represent different orientations. For moderate tilting, an assumed horizontal surface may serve as a local approximation for an actual tilted surface (Larson and Nievinski 2013). Finite plates offer great flexibility in modeling complex surfaces, such as those found in the built environment. A faceted model is more rigorous as edge diffraction is accounted for in addition to ray-tracing reflections. A more specialized surface geometry is that of a spherical Earth, necessary for reflectometry from elevated platforms. Undulated, non-planar surface geometries lead to changes in the reflection power through ray focusing and spreading. Other methods are not ray-based, rather current-based, such as Parabolic Equation (Hannah et al. 1998) and Physical Optics (Chen et al. 2009; Evans et al. 1989; Geren et al. 2008).

Often a total polarization reversal upon reflection is assumed for simplicity, although this is strictly valid only for a perfectly electric conducting (PEC) surface – as if reflections could be suppressed by ensuring that the antenna will respond to RHCP only. In this case the reflection power may be calculated utilizing the LHCP antenna gain pattern and optionally an empirical damping factor. Sometimes this polarization reversal assumption is extended for odd and even number of bounces. A truly polarimetric simulator supports dielectric materials as well. For these cases, the exact reflection polarization ellipticity can be calculated. The appropriate antenna gain patterns can be matched with each co- and cross-polarized reflection components. In contrast to the reflection, the direct/incident signal typically has its polarization ellipticity neglected.

The applicability of each type of simulator depends on the scenario. For example, a geometrical simulation may be adequate when the propagation delay represents the dominating phase component, as is typically the case in tide gauge applications (Larson et al. 2013). A PEC simulator is adequate in predominantly metallic bodies, such as satellites, airplanes, ships, cars, etc. On the other hand, a more complete polarimetric model becomes mandatory in interpreting the retrieval of material compositional properties, such as soil moisture (Zavorotny et al. 2010). A polarimetric simulator is also more realistic for the design of antennas for positioning applications, preventing an otherwise overly-optimistic performance assessment under the assumption of a metallic surface (Chen et al. 2012a; Chen et al. 2012b). Furthermore, polarimetric simulators can support efforts to use polarization diversity to mitigate multipath, as in Groves et al. (2010). Finally, sometimes one is interested not in the deterministic simulation for one specific scattering condition, but in the average and dispersion of an ensemble of such simulations over varying satellite directions and receiver positions (Chen et al. 2010).[1]

Whereas simulators employed for positioning purposes always assume an upright antenna, simulators for coherent reflectometry are typically specialized for tipped or upside-down antennas. This latter design choice simplifies the model as it makes it safer to neglect complementary polarizations (i.e., LHCP or RHCP only), and even the detailed antenna gain pattern under certain symmetry conditions. On the other hand, many reflectometry experiments necessitate more elaborate surface composition simulations, accounting for layering, e.g., air/snow/soil. Another difference is that with research-grade receivers, one would be interested in simulating the full correlation waveform, whereas the use of commercial off-the-shelf receivers requires knowledge of the correlation peak only.

[1] Below we address the statistical averaging of surface scattering over a different domain, random surface roughness, for fixed satellite direction and receiver position.

Finally, users should be aware that atmospheric refraction is not necessarily negligible in reflectometry from elevated platforms. This is the reason why we restrict the applicability of the present model to near-surface installations. Although towers and cliff installations may or may not safely neglect refraction – depending on the vertical separation – we make no claim either way. Further developments as well as a future assessment are warranted.

3.3 FORMULATION

We start by introducing direct and reflected signals, based upon which we define the interferometric and error quantities. The former two signals are more basic or fundamental, while the latter two quantities are of main interest in reflectometry and positioning applications. We proceed to expose the effect of varying coherency on power and phase. Then we examine the code modulation impressed on the carrier wave, with special consideration for multipath or composite-signal reception. We characterize the antenna response in terms of its complex vector effective length, which dictates how the propagating electric field is transformed into a circuit current. The direct and reflected fields are subsequently detailed. The noise power spectral density and bandwidth are defined, and with it the signal to noise ratio. We end with a summary of the expressions developed.

3.3.1 Interferometric and error quantities

Let the direct voltage, collected at the satellite line-of-sight, be $V_\mathrm{d} = |V_\mathrm{d}| \exp(\iota \phi_\mathrm{d})$, with its time-dependent magnitude $|V_\mathrm{d}|$ and phase ϕ_d (with the imaginary unity denoted $\iota = \sqrt{-1}$). The reflection voltage is V_r. Their complex ratio

$$V_\mathrm{i} = V_\mathrm{r}/V_\mathrm{d} \tag{1}$$

is called the interferometric voltage, and it is the main quantity of interest for coherent reflectometry applications. The interferometric phase, $\phi_\mathrm{i} = \phi_\mathrm{r} - \phi_\mathrm{d}$, amounts to the reflection

excess phase with respect to the direct one. The interferometric power, $P_\mathrm{i} = |V_\mathrm{i}|^2 = P_\mathrm{r}/P_\mathrm{d}$, isolates the reflectivity by which the medium responds, independently from the directly collected power. In positioning applications, the reciprocal of interferometric power is known as "D/U", the desired-to-undesired power ratio.

The complex sum of direct and reflection voltages

$$V_\mathrm{c} = V_\mathrm{r} + V_\mathrm{d} = V_\mathrm{d}(1 + V_\mathrm{i}) = V_\mathrm{d}V_\mathrm{e} \tag{2}$$

is the composite voltage. Composite phase, ϕ_c, just like the reflection phase ϕ_r, is contaminated by the direct phase, ϕ_d, which contains many unknown terms (clocks, atmospheric delays, etc.); this makes it difficult to use them for reflection studies. The ratio of composite to direct voltages,

$$V_\mathrm{e} = V_\mathrm{c}/V_\mathrm{d} = 1 + V_\mathrm{i} \tag{3}$$

will be called the error voltage. It is the main quantity of interest for positioning applications because its phase $\phi_\mathrm{e} = \phi_\mathrm{c} - \phi_\mathrm{d}$ quantifies how much phase tracking is in error compared to the assumption of reflection-free or direct-only conditions, i.e., $\phi_\mathrm{c} = \phi_\mathrm{d} + \phi_\mathrm{e}$.

Interferometric and error phases are reckoned from the direct phase, and they can be related in terms of the two-argument four-quadrant arc-tangent:

$$\phi_\mathrm{e} = \mathrm{atan}\big(\sqrt{P_\mathrm{i}}\sin\phi_\mathrm{i}\,, 1 + \sqrt{P_\mathrm{i}}\cos\phi_\mathrm{i}\big); \tag{4}$$

or, assuming $P_\mathrm{r} \ll P_\mathrm{d}$,

$$\phi_\mathrm{e} \approx \sqrt{P_\mathrm{i}}\sin\phi_\mathrm{i} = \mathrm{Im}\{V_\mathrm{i}\} \tag{5}$$

(in radians). So the error phase ϕ_e would seem to be more difficult to model accurately than the interferometric phase ϕ_i, because the former depends additionally on the interferometric power P_i. The corresponding (coherent) powers are related as

$$P_\mathrm{e} = 1 + P_\mathrm{i} + 2\sqrt{P_\mathrm{i}}\cos\phi_\mathrm{i}, \tag{6}$$

$$P_c = P_d + P_r + 2\sqrt{P_d}\sqrt{P_r}\cos\phi_i. \tag{7}$$

In contrast to direct, reflected, and interferometric powers, the error power P_e and the composite power $P_c = P_d P_e$ both include a trigonometric term.

This same nomenclature will be used for the propagation delays as well: τ_d (direct), τ_r (reflected), $\tau_i = \tau_r - \tau_d$ (interferometric), $\tau_c = \tau_d + \tau_e$ (composite), and τ_e (error). There is no such a thing as a composite propagation path, so the composite delay cannot be defined in terms of ray properties; rather, it is to be interpreted as the delay by which a signal replica needs to be shifted, such that is maximizes correlation with the measured composite signal (see below for details). The delay multipath error in general depends on the particular code discriminator employed by the receiver (several are implemented in our simulator), although approximations exist to circumvent it; details will be given below.

3.3.2 Coherence

Composite power $P_c = |V_c|^2 = |V_d + V_r|^2$ should be considered in the average sense:

$$\langle P_c\rangle = \langle V_d^* V_d\rangle + \langle V_r^* V_r\rangle + \langle V_d^* V_r\rangle + \langle V_r^* V_d\rangle, \tag{8}$$

where * denotes complex conjugation and $\langle\cdot\rangle$ is statistical expectation. It is convenient to express it in terms of the complex-valued coherence, $\gamma \equiv \langle V_d^* V_r\rangle / \sqrt{\langle V_d^* V_d\rangle\langle V_r^* V_r\rangle}$:

$$\langle P_c\rangle = \langle P_d\rangle + \langle P_r\rangle + 2\sqrt{\langle P_d\rangle}\sqrt{\langle P_r\rangle}\,|\gamma|\cos\phi_\gamma. \tag{9}$$

The polar decomposition, $\gamma = |\gamma|\exp(\iota\phi_\gamma)$, is especially insightful. Assuming the direct signal to be deterministic causes it to cancel out of the coherence power (squared degree of coherence),

$$|\gamma|^2 = |\langle V_r\rangle|^2 / \langle |V_r|^2\rangle, \tag{10}$$

which can then be interpreted as a measure of the reflection phase purity or variance; the deterministic direct signal remains only in the coherence argument,

$$\phi_\gamma = \arg(\langle V_\mathrm{r}\rangle) - \arg(\langle V_\mathrm{d}\rangle) \tag{11}$$

which is nothing but the interferometric phase, ϕ_i.

Further decomposing the reflection voltage into two components, $V_\mathrm{r} = V_\mathrm{r}^\mathrm{I} + V_\mathrm{r}^\mathrm{C}$, the incoherent component is defined such that its complex product $V_\mathrm{r}^\mathrm{I} V$ is zero on average for voltages other than itself, $V \neq V_\mathrm{r}^\mathrm{I}$, because its phase $\arg(V_\mathrm{r}^\mathrm{I})$ is random (yet its average power $\langle P_\mathrm{r}^\mathrm{I}\rangle$ remains nonzero). In contrast, the coherent component has its average power equal simply to the power implied by its average voltage, $\langle |V_\mathrm{r}^\mathrm{C}|^2\rangle = |\langle V_\mathrm{r}^\mathrm{C}\rangle|^2 = \langle P_\mathrm{r}^\mathrm{C}\rangle$. Thus only the coherent component survives in the average reflection voltage, $\langle V_\mathrm{r}\rangle = \langle V_\mathrm{r}^\mathrm{C}\rangle$, but both components are present in the average reflection power, $\langle P_\mathrm{r}\rangle = \langle P_\mathrm{r}^\mathrm{I}\rangle + \langle P_\mathrm{r}^\mathrm{C}\rangle$.

The coherence phase $\phi_\gamma = \arg(\langle V_\mathrm{r}^\mathrm{C}\rangle) - \arg(\langle V_\mathrm{d}\rangle)$ now involves only the coherent component, and the degree of coherence (squared),

$$|\gamma|^2 = \langle P_\mathrm{r}^\mathrm{C}\rangle / (\langle P_\mathrm{r}^\mathrm{C}\rangle + \langle P_\mathrm{r}^\mathrm{I}\rangle), \tag{12}$$

reveals to be simply the coherent fraction of the total reflected power; a binomial expansion $|\gamma| \approx 1 - 0.5\langle P_\mathrm{r}^\mathrm{I}\rangle / \langle P_\mathrm{r}^\mathrm{C}\rangle$ shows further that it decreases with increasing incoherent-to-coherent power ratio (assuming $\langle P_\mathrm{r}^\mathrm{C}\rangle \gg \langle P_\mathrm{r}^\mathrm{I}\rangle$). The matching between the scattered power spectral distribution on the one hand and the measurement averaging period on the other hand will dictate what proportion of the total reflected power can be captured coherently – essentially a low-pass filter. For the same physical scattering process, the amount of coherently collected power can be varied employing varying coherent integration periods. The degree of coherence quantifies this gradation in a continuum between 0 and 1.

Substituting the coherent and incoherent powers into the composite power, and dropping the brackets notation, we finally obtain:

$$P_{\mathrm{c}} = P_{\mathrm{d}} + P_{\mathrm{r}}^{\mathrm{C}} + 2\sqrt{P_{\mathrm{d}}}\sqrt{P_{\mathrm{r}}^{\mathrm{C}}}\cos\phi_{\mathrm{i}} + P_{\mathrm{r}}^{\mathrm{I}} = \left|V_{\mathrm{d}} + V_{\mathrm{r}}^{\mathrm{C}}\right|^2 + P_{\mathrm{r}}^{\mathrm{I}}. \quad (13)$$

The degree of coherence $|\gamma|$ disappears and we are left with the interference of direct and coherently reflected voltages, in addition to the incoherent reflected power.

3.3.2.1 Incoherent power

Incoherent reflections are incapable of affecting the expected value of the interferometric phase, only its variance – i.e., as the magnitude of the complex coherence γ diminishes, it becomes more difficult to recover its phase. In forward modeling GPS multipath observables, neglecting incoherent power $P_{\mathrm{r}}^{\mathrm{I}}$ is inconsequential for carrier phase. SNR is affected, but only in its trend, $\mathrm{tSNR} \propto P_{\mathrm{d}} + P_{\mathrm{r}}^{\mathrm{I}} + P_{\mathrm{r}}^{\mathrm{C}}$ – which is typically a monotone function of elevation angle – over which the interference fringes $\mathrm{dSNR} \propto 2\sqrt{P_{\mathrm{d}}}\sqrt{P_{\mathrm{r}}^{\mathrm{C}}}\cos\phi_{\mathrm{i}}$ are superimposed. Finally, pseudorange can be influenced by $P_{\mathrm{r}}^{\mathrm{I}}$, depending on the code discriminator employed. We will neglect incoherent power from now on; we will model the coherent reflection only, $V_{\mathrm{r}} = V_{\mathrm{r}}^{\mathrm{C}}$, whose magnitude and phase will incorporate the effects of varying coherence.

3.3.3 Code modulation

The voltages V above are the result of a matching filter that correlates the received voltage Z against a replica Z_0 of the transmitted signal over a given coherent integration time T:

$$V = \frac{1}{T}\int_{-T/2}^{+T/2} Z^* Z_0(t)\,\mathrm{d}t \quad (14)$$

The replica $Z_0(t) = C(t-\tau_0)\exp(\iota 2\pi f_0 t)$ mimics the code modulation C and carrier frequency f impressed on received voltage $Z = Y\,C(\tau)\exp(\iota 2\pi f t)$. The post-correlation result can be expressed as $V = Y\,W\,\Psi$, where the pre-correlation voltage Y is described below. The unity complex factor $\Psi = \exp\big(\iota 2\pi(\delta\tau + \delta f T)\big)$ accounts for an out-of-lock phase change. The real-valued Woodward ambiguity function $W \approx \Lambda\,\Gamma$ can be separated into two factors, a product of

the code auto-correlation $\Lambda(\delta\tau)$ (function of delay difference $\delta\tau = \tau_0 - \tau$ and the code chipping rate) and a normalized sinc function $\Gamma(\delta f) = \mathrm{sinc}(\delta f T) = \sin(\pi\delta f T)/(\pi\delta f T)$ (dependent on the frequency difference $\delta f = f_0 - f$ accumulated coherently during T). This separability applies to BPSK modulations (Zavorotny and Voronovich 2000), as utilized in all legacy GPS signals, as well as in the newer L2C signal and the newest L5 signal; it will not fully hold for the future GPS L1C signal, whose design is based on a BOC modulation.

In the case of multipath reception, the matching is done using a single replica against the composite voltage. In this case the replica is locked such that $\Psi = 1$. The direct and reflection delay tracking differences are, respectively,

$$\delta\tau_\mathrm{d} = \tau_0 - \tau_\mathrm{d} = \tau_\mathrm{e}, \tag{15}$$

$$\delta\tau_\mathrm{r} = \tau_0 - \tau_\mathrm{r} = \tau_\mathrm{e} + \tau_\mathrm{d} - \tau_\mathrm{r} = \tau_\mathrm{e} - \tau_\mathrm{i}. \tag{16}$$

The corresponding frequency differences are similar, $\delta f_\mathrm{d} = \Delta f_\mathrm{e}$, $\delta f_\mathrm{r} = \delta f_\mathrm{e} - \Delta f_\mathrm{i}$, and in fact are just delay-rates scaled by wavelength, e.g., $\Delta f_\mathrm{i} = \dot{\tau}_\mathrm{i}/\lambda = \Delta f_\mathrm{r} - \Delta f_\mathrm{d}$, where the dot denotes time-derivative. These frequencies are not to be confused with the Doppler shifts experienced by the direct and reflected signals, $\Delta f_\mathrm{d} = \dot{\tau}_\mathrm{d}/\lambda$, $\Delta f_\mathrm{r} = \dot{\tau}_\mathrm{r}/\lambda$, whose absolute value is generally much greater than δf_d, δf_r. The direct and reflected ambiguity functions thus read $W_\mathrm{d} = \Lambda(\tau_\mathrm{e})\,\Gamma(\Delta f_\mathrm{e})$ and $W_\mathrm{r} = \Lambda(\tau_\mathrm{e} - \tau_\mathrm{i})\,\Gamma(\Delta f_\mathrm{e} - \Delta f_\mathrm{i})$. Everything else being the same, stationary multipath is more severe than fast-changing multipath, because of the suppression offered by the sinc function Γ.

Given the interferometric delay τ_i and pre-correlation direct and reflection voltages Y_d, Y_r, the delay error τ_e can be calculated rigorously employing a code discriminator function against the composite voltage $V_\mathrm{c} = V_\mathrm{d} + V_\mathrm{r} = Y_\mathrm{d}W_\mathrm{d} + Y_\mathrm{r}W_\mathrm{r}$. An initial guess $\tau_\mathrm{e} = 0$ is improved iteratively until convergence, applying the corrections provided by the discriminator. This is akin to what a receiver performs in real-time. The Doppler error Δf_e could be obtained via numerical

differentiation of a time-succession of so-obtained delay errors τ_e. Notice the interferometric Doppler Δf_i is non-zero even for a stationary receiver and stationary surface, except when the satellite direction is also stationary – it is a consequence of the displacement of the specular point. We will ignore Doppler from now on.

For small interferometric delays, it may be acceptable to neglect code modulation entirely, as the approximation $Y_\mathrm{c} \approx V_\mathrm{d} + V_\mathrm{r}$ is reasonably accurate in terms of the phase error ϕ_e and also composite power P_c. Moreover, in this case the error delay can be approximated as:

$$\tau_\mathrm{e} \approx \tau_\mathrm{i}\sqrt{P_\mathrm{i}}\cos\phi_\mathrm{i}\,/\big(1+\sqrt{P_\mathrm{i}}\cos\phi_\mathrm{i}\big). \tag{17}$$

Compounding this small-delay approximation with the previous small-power approximation ($P_\mathrm{i} \ll 1$) we obtain:

$$\tau_\mathrm{e} \approx \tau_\mathrm{i}\,\mathrm{Re}\{V_\mathrm{i}\} = \tau_\mathrm{i}\sqrt{(1-\phi_\mathrm{e}^2)}. \tag{18}$$

Later we will assess the accuracy of these expressions. Such insensitivity of the error delay to any particular code discriminator is helpful for reflectometry applications because environmental retrievals are less likely to be receiver-dependent. This also implies that all receivers are equally bad in mitigating short-delay multipath, which is unfortunate for positioning applications. It is only for large interferometric delays that the code modulation becomes effective in suppressing the reflection voltage V_r, which contributes less and less compared to the direct one V_d, eventually being rejected from the composite voltage V_c as τ_i exceeds a threshold (beyond the code chip width).

3.3.4 Antenna response

The pre-correlation voltages Y are scalars that result from the dot-product of a vector-valued electric field $\bar{E} = [E^\mathrm{R},\ \ E^\mathrm{L}]^\mathrm{T}$ (in terms of right- and left-handed polarization components –

RHCP and LHCP) against the antenna complex vector effective length (in meters), thus converting from volts-per-meter to volts (Milligan 2005):

$$Y_{\mathrm{d}} = \bar{L}_{\mathrm{d}}^{\dagger} \bar{E}_{\mathrm{d}} \tag{19}$$

$$Y_{\mathrm{r}} = \bar{L}_{\mathrm{r}}^{\dagger} \bar{E}_{\mathrm{r}} \tag{20}$$

where the dagger † denotes conjugate transposition. The subscript serves as shorthand for the direction at which the antenna response is evaluated, e.g., for $\bar{L}_{\mathrm{d}}$, the direct signal's boresight angle and axial angle in the antenna body-fixed coordinate system.

The vector norm $L = \|\bar{L}\| = (|L^{\mathrm{R}}|^2 + |L^{\mathrm{L}}|^2)^{1/2}$ is given by (Milligan 2005):

$$L = L_{\mathrm{iso}} \sqrt{G}. \tag{21}$$

The complex effective length of an isotropic antenna is $L_{\mathrm{iso}} = 2\sqrt{(Z/Z_0)\lambda^2/(4\pi)}$, and the effective area of an isotropic antenna, $\lambda^2/(4\pi)$; Z is the load impedance (in ohms) and Z_0 is the vacuum impedance. The antenna gain G (usually given in decibels, $G = 10 \log_{10} G_{\mathrm{dB}}$) – not to be confused with the peak gain – is direction-dependent but polarization-independent. The polarization dependency is represented by the complex unity vector $\hat{L} = \bar{L}/L$:

$$\hat{L} = [\hat{L}^{\mathrm{R}}, \quad \hat{L}^{\mathrm{L}}]^{\mathrm{T}} = [\sqrt{G^{\mathrm{R}}} \exp(\iota \Phi^{\mathrm{R}}), \quad \sqrt{G^{\mathrm{L}}} \exp(\iota \Phi^{\mathrm{L}})]^{\mathrm{T}} / \sqrt{G}. \tag{22}$$

The magnitudes follow from the respective partial power gains for each RHCP and LHCP (G^{R}, G^{L}), and similarly for the total power gain, $G = G^{\mathrm{R}} + G^{\mathrm{L}}$; the respective antenna phase patterns are denoted $\Phi^{\mathrm{R}}, \Phi^{\mathrm{L}}$. Defining the antenna polarimetric power ratio, $G^{\mathrm{L/R}} = G^{\mathrm{L}}/G^{\mathrm{R}}$, as well as the antenna polarimetric phase difference, $\Phi^{\mathrm{L-R}} = \Phi^{\mathrm{L}} - \Phi^{\mathrm{R}}$, we can write the antenna complex vector effective length as:

$$\bar{L} = [L^{\mathrm{R}}, \quad L^{\mathrm{L}}]^{\mathrm{T}} = L_{\mathrm{iso}} \sqrt{G^{\mathrm{R}}} \exp\left(-\iota \Phi^{\mathrm{R}}\right) [1, \quad \sqrt{G^{\mathrm{L/R}}} \exp(-\iota \Phi^{\mathrm{L-R}})]^{\mathrm{T}}. \tag{23}$$

(Antenna phases are denoted uppercase, e.g., $\Phi_{\mathrm{d}}^{\mathrm{R}} = \arg\left(L_{\mathrm{d}}^{\mathrm{R}}\right)$, and electric field phases are lowercase, e.g., $\phi_{\mathrm{d}}^{\mathrm{L}} = \arg\left(E_{\mathrm{d}}^{\mathrm{L}}\right)$; voltage phases are also lowercase but need no superscript, e.g., $\phi_{\mathrm{r}} = \arg(V_{\mathrm{r}})$, as they have no polarization.)

3.3.4.1 Antenna gain

The antenna gain pattern, G, for each polarization and as evaluated in each direct and reflected direction $G_{\mathrm{d,r}}^{\mathrm{R,L}}$, can be measured in anechoic chambers. GPS antennas are typically set upright, being omni-directional in azimuth and hemispherical in elevation angle, to allow multiple satellites to be tracked simultaneously while minimizing ground noise reception. Sometimes the antenna is turned upside-down, or tipped with its boresight facing the horizon; see sec. 3.7 for discussion. By transforming the viewing direction – from east, north, up coordinates to antenna body-aligned components, arbitrary antenna orientations are reduced to the upright installation case.

The antenna gain pattern is normally made available as a principal plane cut, vertically across the antenna axis. This offers some information about the axial asymmetry in the antenna horizontal plane. Given irregularly-spaced gain samples, we fit a set of spherical harmonics (only zonal and first-order tesseral harmonics). The resulting coefficients can then be evaluated to obtain the antenna gain over a regularly spaced grid.

3.3.4.2 Antenna phase

The receiver antenna phase pattern is typically only known in RHCP, Φ^{R}, because this has the greatest impact on positioning applications. It can be modeled as

$$\Phi^{\mathrm{R}} = -\vec{\Phi}^{\mathrm{R}} \cdot \hat{\rho} + \tilde{\Phi}^{\mathrm{R}}. \tag{24}$$

The first term consists of the phase center offset, $\vec{\Phi}^{\mathrm{R}} = \left[\Phi_{\mathrm{x}}^{\mathrm{R}}, \Phi_{\mathrm{y}}^{\mathrm{R}}, \Phi_{\mathrm{z}}^{\mathrm{R}}\right]^{\mathrm{T}}$ (a spatial vector, $\vec{\Phi} \in \mathbb{R}^3$, not to be confused with EM vectors, $\bar{E} \in \mathbb{C}^2$), projected on the viewing direction, $\hat{\rho}$. The second term $\widetilde{\Phi}^{\mathrm{R}}$ is the phase center variation, a scalar-valued function describing the asphericity of the wavefronts generated by an astigmatic antenna; it is millimetric near the boresight for geodetic-quality antennas. Direction-dependent quantities can be evaluated at the reflection or satellite direction, i.e., corresponding to a given azimuth and elevation angle, e.g., $\Phi_{\mathrm{d}}^{\mathrm{R}}, \hat{\rho}_{\mathrm{r}}, \widetilde{\Phi}_{\mathrm{r,d}}^{\mathrm{R}}$; in contrast, $\vec{\Phi}^{\mathrm{R}}$ is a constant.

The total pattern tends to be dominated by a cm-level vertical component of the offset vector (assuming an upright installation), $\Phi^{\mathrm{R}} \approx -\Phi_{\mathrm{z}}^{\mathrm{R}} \sin \mathrm{e}$. This yields a phase excess and deficit at zenith and nadir, respectively; the effect is similar to an antenna being at a different height. Under the assumption of a planar horizontal surface, the reflection elevation angle is trivially $\mathrm{e}_{\mathrm{r}} = -\mathrm{e}_{\mathrm{d}}$. The antenna phase difference, $\Phi^{\mathrm{L-R}} = \Phi^{\mathrm{L}} - \Phi^{\mathrm{R}}$, evaluated across different polarizations but at the same viewing direction is less well-known and approximated here as -90°.

3.3.5 Direct electric field

The direct electric field is expressed as:

$$\bar{E}_{\mathrm{d}} = [E_{\mathrm{d}}^{\mathrm{R}}, \quad E_{\mathrm{d}}^{\mathrm{L}}]^{\mathrm{T}} = \left(\sqrt{P_{\mathrm{d}}^{\mathrm{R}}}/L_{\mathrm{iso}}\right) \exp(\iota \phi_{\mathrm{d}}^{\mathrm{R}}) \left[1, \quad E_{\mathrm{d}}^{\mathrm{L/R}}\right]^{\mathrm{T}} \tag{25}$$

where $P_{\mathrm{d}}^{\mathrm{R}}/L_{\mathrm{iso}}^2$ (units $\mathrm{W/m^2}$) is the power spatial density. The direct polarimetric field ratio is $E_{\mathrm{d}}^{\mathrm{L/R}} = E_{\mathrm{d}}^{\mathrm{L}}/E_{\mathrm{d}}^{\mathrm{R}} = \sqrt{P_{\mathrm{d}}^{\mathrm{L/R}}} \exp(\iota \phi_{\mathrm{d}}^{\mathrm{L-R}})$. The direct polarimetric power ratio,

$$P_{\mathrm{d}}^{\mathrm{L/R}} = \left(\frac{\alpha_{\mathrm{d}} - 1}{\alpha_{\mathrm{d}} + 1}\right)^2 \tag{26}$$

is related to the direct field polarization ellipticity:

$$\alpha_{\mathrm{d}} = \frac{\left|E_{\mathrm{d}}^{\mathrm{R}}\right| + \left|E_{\mathrm{d}}^{\mathrm{L}}\right|}{\left|E_{\mathrm{d}}^{\mathrm{R}}\right| - \left|E_{\mathrm{d}}^{\mathrm{L}}\right|}. \tag{27}$$

The signals broadcast by GPS satellites are predominantly RHCP, with the LHCP magnitude specified never to exceed 20% of the RHCP (ellipticity $\alpha_{\mathrm{d}} \leq$ 3.2 dB [GPSW, 2010]). Under the assumption $P_{\mathrm{d}}^{\mathrm{L/R}} = 0$, the direct polarimetric phase difference $\phi_{\mathrm{d}}^{\mathrm{L-R}} = \phi_{\mathrm{d}}^{\mathrm{L}} - \phi_{\mathrm{d}}^{\mathrm{R}}$ becomes irrelevant, which is fortunate, as this value is currently unknown (if it becomes known, it can be inserted in the simulator).

3.3.5.1 Transmitted signal

The various GPS frequency/code signal combinations have different specified received power levels, which shift the SNR decibel curves up and down. The satellite antenna gain patterns are not made available to the public, or else one could easily compute the expected power level based on the ephemeris-calculated satellite-receiver distance. In the lack of such ancillary information, we rely on the fact that satellite gain patterns were designed to compensate for the increased range, from satellite nadir to Earth's limb (or from receiver's zenith to horizon), keeping the variations in received power level to within 2 dB. The power available for an isotropic antenna in GPS is typically $P_{\mathrm{d}}^{\mathrm{R}} \sim -160$ dB-W. More exact values as well as the remaining variation are specified in GPSD-USAF (2011) as a function of elevation angle and are incorporated in model.

3.3.6 Reflected electric field

The reflected electric field is decomposed as $\bar{E}_{\mathrm{r}} = SI\bar{\bar{R}}\bar{E}_{\mathrm{d}}$. It is expressed in terms of the direct field $\bar{E}_{\mathrm{d}}$, incident on the receiving antenna, not the field incident on the surface. In the following we explain the remaining components, in the order that they are applied to $\bar{E}_{\mathrm{d}}$.

3.3.6.1 Medium composition

The reflection matrix $\bar{\bar{R}} \in \mathbb{C}^{2\times 2}$,

$$\bar{\bar{R}} = R^{\mathrm{S}}\bar{\bar{P}}^{\mathrm{S}} + R^{\mathrm{X}}\bar{\bar{P}}^{\mathrm{X}} \tag{28}$$

is a combination of same- and cross-sense polarizing matrices, where $\bar{\bar{P}}^{\mathrm{S}}$ is the 2-by-2 identity matrix and $\bar{\bar{P}}^{\mathrm{X}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is the first Pauli matrix. The circularly polarized scalar reflection coefficients are defined as:

$$R^{\mathrm{S}} = (R^{\mathrm{H}} + R^{\mathrm{V}})/2, \tag{29}$$

$$R^{\mathrm{X}} = (R^{\mathrm{H}} - R^{\mathrm{V}})/2. \tag{30}$$

The linearly polarized reflection coefficients follow from the solution of the Fresnel equations for two homogeneous half-spaces:

$$R^{\mathrm{H}} = (\epsilon \cos\theta - \xi)/(\epsilon \cos\theta + \xi) \tag{31}$$

$$R^{\mathrm{V}} = (\cos\theta - \xi)/(\cos\theta + \xi) \tag{32}$$

where $\xi = \sqrt{(\epsilon - \sin^2\theta)}$; the angle of incidence (with respect to the surface normal) is denoted θ; the permittivity ratio, $\epsilon = \epsilon_{\mathrm{b}}/\epsilon_{\mathrm{t}}$, relates bottom and top halfspaces. The top one ϵ_{t} can typically be assumed unity, for air, except when dealing with layered media. The bottom permittivity $\epsilon_{\mathrm{b}} = \epsilon_{\mathrm{b}}' + \iota\epsilon_{\mathrm{b}}''$ is made of real and imaginary components, the latter of which can also be related to conductivity σ as:

$$\epsilon'' = \sigma/(2\pi f \epsilon_0) \approx \sigma\, \lambda\, 60\ \mathrm{S}^{-1}, \tag{33}$$

in terms of the carrier frequency f (in hertz); wavelength λ; a derived constant $60\ \mathrm{S}^{-1} \approx \mu_0 c_0/(2\pi)$, in units of reciprocal of Siemens; and vacuum constants: permittivity ϵ_0, permeability μ_0, and speed of light, c_0.

3.3.6.2 Interface geometry

The complex-valued scalar $I = |I| \exp(\iota\phi_I)$ has phase $\phi_I = k\tau_\mathrm{i}$, where $k = 2\pi/\lambda$ is the wavenumber and τ_i is the interferometric delay. For a planar horizontal surface, the latter is simply (Georgiadou and Kleusberg 1988):

$$\tau_\mathrm{i} = 2H \sin e_\mathrm{d}, \tag{34}$$

in terms of the reflector height (height of the antenna above the ground), H, and the direct signal elevation angle e_d (with respect to the receiver local horizon). The interferometric delay-rate, assuming a constant reflector height, is:

$$\dot{\tau}_\mathrm{i} = 2H\dot{e}_\mathrm{d} \cos e_\mathrm{d}. \tag{35}$$

The magnitude is $|I| = 1$, which means that the additional propagation distance τ_i does not create a free-space propagation loss; this follows from the assumption of a plane incident wave. The sign of the propagation phase must be consistent with the sign of the imaginary component of permittivity, such that forward propagation in a lossy medium ($\Im\{\epsilon\} > 0$) does indeed lead to power loss: $\lim_{\tau\to\infty}\left|\exp(\iota k\tau\sqrt{\epsilon})\right| = 0$.

3.3.6.3 Random surface roughness

The magnitude $|S|$ represents a loss of coherent power. We calculate it from the theories of coherent scattering from random surfaces (Beckmann and Spizzichino 1963), as:

$$|S| = \exp(-0.5k^2 s_h^2 \cos^2\theta), \tag{36}$$

where s_h is the surface height standard deviation (in meters); notice it is polarization-independent. Random roughness has been reported to also have an effect on phase, ϕ_S, causing an apparent surface raise (diminished reflector height); this follows from the preferential illumination of surface crests, compared to shadowed surface troughs (Bourlier et al. 2006). Notice the power effect is greatest at normal incidence, whereas the phase effect would be

greatest at grazing incidence. Currently the forward model only accounts for the magnitude effect, thus assuming $S = |S|$. Bourlier et al. (2006) state that the shadowing effect can be neglected for grazing angles greater than $\mathrm{atan}\left(2\sqrt{2}\, s_{h'}\right)$, where $s_{h'}$ is the standard deviation of surface slopes; users should be aware of this caveat, especially those dealing with dynamic surfaces, as in ocean scattering.

3.3.7 Noise power

In the GPS literature it is customary to introduce the notation C for effective carrier power, which includes receiver losses absent in P_c as above; typically these are power transmission loss between the antenna and the preamplifier (Spilker et al. 1996). Here we neglect that, assuming $C = P_\mathrm{c}$. Carrier power becomes combined with the noise power spectral density N_0 to form the carrier-to-noise-density ratio, C/N_0 (in hertz); and with the noise bandwidth, B_n, to form the carrier-to-noise ratio, C/P_n (in watts per watt) via the noise power $P_\mathrm{n} = N_0 B_\mathrm{n}$. The RINEX "S1" and "S2" observables will be taken here as $\mathrm{SNR} = C/P'_\mathrm{n}$, in terms of a modified P'_n, defined below.

Noise density $N_0 = k_\mathrm{B} T$ is calculated as the product of the Boltzmann constant, $k_\mathrm{B} \approx 1.38 \times 10^{-23}$ J/K, and a noise-equivalent temperature, $T = T_\mathrm{A} + T_\mathrm{R}$. The antenna contribution T_A lies in the range $75 - 130$ K for a typical installation and hemispherical gain (Langley 1997), and reaches a much higher ambient temperature ($\sim$290 K) when the antenna is replaced by a hardware simulator. The receiver contribution is taken as $T_\mathrm{R} \approx 470$ K, corresponding to a circuit made of a low-noise amplifier in-between a short and long cables, connecting the antenna element to the receiver (Misra and Enge 2006). Both values depend on the installation, i.e. whether the antenna is upright, the length and physical temperature of the cables, etc. The noise

density is the same for all satellites tracked at the same time with the same antenna, but it does vary over time.

We define a modified noise power as $P_{\mathrm{n}}' = K' P_{\mathrm{n}}$. The non-negative unitless factor K' represents tracking losses, mainly related to the code modulation, which can be quite severe and disfigure the multipath modulation otherwise clearly present in SNR. These are, for example, codeless tracking losses affecting the encrypted P(Y) codes (Woo 2000) and cross-channel self-interference (i.e., between different satellites), which significantly impact the shorter C/A codes (Lestarquit and Nouvel 2012). For the newer L2C code, K' can be neglected without degradation in the model/observation agreement; this is because these codes, contrary to the P(Y), are publicly known; and in contrast to C/A, are sufficiently long (Fontana et al. 2001).

3.3.8 Simplified expressions

The rigorous forward model exposed above relies on matrix/vector operations. Alternatively, we can write the resulting direct and reflected powers in full as (notice the isotropic antenna complex effective length L_{iso} cancels out):

$$P_{\mathrm{d}} = P_{\mathrm{d}}^{\mathrm{R}} G_{\mathrm{d}}^{\mathrm{R}} W_{\mathrm{d}}^2, \tag{37}$$

$$P_{\mathrm{r}} = P_{\mathrm{d}}^{\mathrm{R}} |XSW_{\mathrm{r}}|^2, \tag{38}$$

where a purely RHCP incident signal was assumed. This assumption – that can be disabled in the simulator, but without which the mathematical formulae quickly become unyielding – allows us to interpret the like-polarization reflection coefficient as RHCP-producing, and similarly for LHCP/cross-polarization. We can further define coupled surface/antenna coefficients,

$$X^{\mathrm{R}} = R^{\mathrm{S}} \sqrt{G_{\mathrm{r}}^{\mathrm{R}}} \exp(\iota \Phi_{\mathrm{r}}^{\mathrm{R}}), \tag{39}$$

$$X^{\mathrm{L}} = R^{\mathrm{X}} \sqrt{G_{\mathrm{r}}^{\mathrm{L}}} \exp(\iota \Phi_{\mathrm{r}}^{\mathrm{L}}), \tag{40}$$

as well as their complex sum, $X = X^{\mathrm{R}} + X^{\mathrm{L}}$. The interferometric phase then reads:

$$\phi_{\mathrm{i}} = \phi_X + \phi_I - \Phi_{\mathrm{d}}^{\mathrm{R}}. \tag{41}$$

The first term $\phi_X = \arg(X)$ accounts for the surface and antenna effects on the reflection; the second one $\phi_I = k\tau_{\mathrm{i}}$ follows from the interferometric propagation delay; the third one $\Phi_{\mathrm{d}}^{\mathrm{R}}$ is the antenna phase contribution to the direct signal. These are the terms not in common between the direct and reflected voltage phases:

$$\phi_{\mathrm{d}} = \phi_{\mathrm{d}}^{\mathrm{R}} + \Phi_{\mathrm{d}}^{\mathrm{R}}, \tag{42}$$

$$\phi_{\mathrm{r}} = \phi_{\mathrm{d}}^{\mathrm{R}} + \phi_X + \phi_I. \tag{43}$$

Notice the direct electric field phase $\phi_{\mathrm{d}}^{\mathrm{R}}$ cancels out of the interferometric phase $\phi_{\mathrm{i}} = \phi_{\mathrm{r}} - \phi_{\mathrm{d}}$. The fact that we can safely ignore $\phi_{\mathrm{d}}^{\mathrm{R}}$ in multipath modeling is tremendously useful, because we need not consider clock errors, atmospheric propagation delays – at least the bulk of it, especially for near-surface antennas – and other effects that would otherwise need to be accounted for in a positioning solution.

3.4 SIMULATIONS

We now illustrate the main features of the forward model. We begin with simulations for a typical scenario found at geodetic installations. We proceed to explore the various GPS transmitted signals. It is followed by an exploration of the surface characteristics. We finish up considering the impact of the antenna setup – its height above the ground, orientation, and gain pattern. The implications for positioning and reflectrometry applications are emphasized.

3.4.1 Nominal conditions

We assume a horizontal ground surface, made up of soil with medium-level moisture and negligible roughness, observed with a choke-ring antenna installed upright on a 1.5-m tall monument. Figure 5 shows the model results for the L2C signal observables: SNR, carrier-phase multipath error, and pseudorange code multipath error, where error means the difference with

respect to multipath-free conditions. As one would expect, carrier-phase multipath error is at least an order of magnitude smaller than the pseudorange multipath error and is limited to a quarter of wavelength provided the magnitude of the direct voltage remains greater than the reflected one (this condition can be violated in the presence of multiple reflections). Besides the rigorous solution, in the same figure we show also results from the small-magnitude and small-delay approximations, eq. (5), (17), (18).

All three observables exhibit a series of peaks and troughs, or fringes. Noticeable features in the interference patterns are the spacing between fringes, the horizontal position of the pattern as a whole, and magnitude variations. Although none of the observables is perfectly sinusoidal, a sinusoid can still be fit after some manipulation of the data. This fitting serves to quantify the multipath modulation frequency, phase-shift, and amplitude. In a variety of cases, the best-fitting sinusoid parameters can be related to physical properties describing the environmental conditions in the antenna surroundings. For linearly polarized antennas the multipath effect requires a different fitting procedure (Rodriguez-Alvarez et al. 2011b).

Figure 6 shows the constituting quantities responsible for producing the observables presented in Figure 5. The top panel demonstrates how interferometric power P_{i} is direction-dependent, increasing with decreasing elevation angle, as the direct and reflected power converge. This violation in the common assumption $P_{\mathrm{r}} \ll P_{\mathrm{d}}$ follows from the surface/antenna polarization matching, as detailed below. The bottom panel separates the geometrical ϕ_I and compositional ϕ_X components of the interferometric phase ϕ_{i}, that depend on the propagation delay and surface material, respectively. The former is a linear function of sine of elevation angle, $\tau_{\mathrm{i}} = 2\,H \sin e$, whereas the latter is a more complicated sigmoid-like function (it would be essentially a constant for a metallic surface). It should be highlighted that the complicated

oscillations present in error power and phase follow from simpler monotone variations in the underlying interferometric quantities. Consequently, interferometric parameters could be more tightly constrained than error parameters if using measurements to model multipath.

3.4.2 Code modulation

An important characteristic of the transmitted signal is the code modulation impressed on the carrier, for ranging as well as frequency-band sharing. Signals currently broadcast by most GPS satellites are L1-C/A, L1-P(Y), L2-P(Y), and L2-C (or simply L2C; Fontana et al. 2001). The most important aspect of the code modulation for multipath modeling is the chipping rate, which for P(Y) is ten times higher than for either C/A or L2C. A higher chipping rate has better rejection performance for large interferometric delays because it narrows the code auto-correlation function $\Lambda(\delta\tau)$. Yet both rates remain largely ineffective against small-delay reflections, as observed near grazing incidence.

Another aspect of the code modulation that impacts SNR is the code length. It dictates the amount of cross-channel self-interference. For example, a high-power, high-elevation angle satellite may create spurious correlations in tracking a lower-power rising or setting satellite. The C/A code has shortest length thus it is the most susceptible to this problem. The issue seems to be exacerbated by small Doppler differences between the two satellites; also certain specific PRN are known to be more vulnerable than others; see Lestarquit and Nouvel (2012) and references therein. As these events are caused by the presence of other satellites simultaneously in view, their times of occurrence are largely repeatable – with sidereal day, similar as multipath itself –, yet their magnitude is expected to be less predictable. C/A losses are not currently contemplated in the forward model, so one should expect larger deviations for this particular GPS signal, compared to L2C SNR, which has a longer code.

The P(Y) code, in both L1 and L2 frequencies, requires codeless tracking of the encrypted Y code when using civilian receivers. Woo (2000) reports systematic losses, inversely proportional to C/N_0 itself, which affect primarily the trend tSNR but also the fringes dSNR to some extent. We have incorporated empirical calibration curves developed using simultaneously measured L2-P(Y) and L2C SNR. As this effect is receiver-dependent, this calibration needs be repeated for other models (we used a Trimble NetR8). Compared to C/A, the P(Y) tracking losses have a less random, more predictable character (Figure 7).

3.4.3 Antenna height

Changing the height of the antenna above the ground or reflecting surface changes primarily the modulation frequency and phase shift (Figure 8). This is caused by the interferometric delay. For a horizontal surface, the specular point gets closer to the antenna with decreasing reflector height, yet the incident and reflection directions remain the same. Consequently, there are no changes in the surface/antenna response, which is function of the incident and reflection angles. The only change in the modulation amplitude is caused by the code modulation, which decreases the reflected power with increasing delay, although it would only be significant for very large reflector heights.

The ideal reflector height depends on the purpose of the application. For example, one might situate an antenna differently for reference frame realization vs. real-time deformation monitoring. In the latter application, high-rate displacements are the quantity of interest, so an antenna closer to the ground would introduce fewer artifacts in the position time series. For a reference frame site, absolute position biases are to be avoided, so a taller antenna would seem preferable, because multipath errors would more likely average to zero. Although code pseudorange RMS error increases without bound in proportion to reflector height, carrier phase

observations are more important for long static positioning sessions. For reflectometry, taller antennas are preferred; e.g., for snow sensing, the antenna should be built much taller (more than two wavelengths) than the highest expected snow level.

3.4.4 Random surface roughness

Surface height standard deviation s_h serves as a parameterization for loss of coherent power (recall that we have incorporated coherence into the reflection voltage). Various phenomena other than random surface roughness can cause loss of coherence, from clock dither, to atmospheric turbulence, and volumetric inhomogeneities. Thus care needs to be exercised in interpreting s_h as an effective parameter or equivalent roughness, an amalgamation of different sources affecting coherently reflected power.

Degree of coherence reduction will decrease the visibility of interference fringes. For example, in optical interferometry, intensity fringes become fainter; in radar imaging, phase fringes become noisier. In GPS, SNR multipath modulation amplitude decreases; if we could measure the interferometric phase directly, it would be noisier but its mean value would be unaffected. In contrast, the error phase would predominantly diminish in magnitude because it also involves the (coherent) interferometric power, $\phi_{\mathrm{e}} \approx \sqrt{P_{\mathrm{i}}} \sin \phi_{\mathrm{i}}$.

In this forward model we account explicitly only for loss of coherence due to surface random roughness, which is driven by a single free parameter, surface height standard deviation (with respect to a trend surface, possibly undulated). Notice that height correlation length does not directly affect the coherent power, although the former is assumed to be much smaller than the illuminated portion of the surface (so that several roughness cycles contribute to the reflection). The average of surface deviations, or trend surface, does not suffer randomization

during the coherent integration period, and would require a deterministic rather than stochastic model (not covered here).

Increasing surface roughness decreases the magnitude of the multipath signal in all three GPS observables (Figure 9), and it does so in an elevation angle dependent manner: higher/lower elevation observations are more/less affected by roughness. Consequently, installing a geodetic monument in a site where the antenna is surrounded by randomly shaped objects, e.g. rocks, would be a valid multipath mitigation strategy. The efficiency would depend on the size of objects, although they would not need to be made of radiofrequency-absorbing material, or metal, or have any other specific composition. Vegetation plays a similar role. For the same reason that surface roughness is benign for positioning applications, it represents a fundamental physical limit for coherent reflectometry, as it may extinguish the multipath modulation and thus, the environmental retrieval.

3.4.5 Surface material

Material composition – types (water, concrete, soil, etc.) and their properties (soil moisture, snow density, etc.) – has an impact on all aspects of the multipath modulation signal (Figure 10). The medium is modeled as an effective homogeneous material with an equivalent complex-valued permittivity, which is input into the conventional Fresnel reflection formulae.

Each material produces different GPS observations, not just directly through the Fresnel reflection coefficients themselves, but also because they elicit a different response from the antenna, depending on the reflection polarization. Two polarization regimes are demarcated by the Brewster angle, separating LHEP reflections (left-handed elliptically polarized) at higher elevation angles from RHEP at lower angles. The Brewster angle is lowered by the medium conductivity, e.g., it is about 10 degrees for wet ground versus ~ 25° for dry ground (Figure 11).

Although a metal surface yields a strong reflected electric field, it results in a weak reflected voltage as the LHEP field is captured by a RHEP antenna. Reflections off-dielectric media, such as most natural land surfaces, retain the polarization of the incident vector electric field in the limit of grazing incidence. In this case the antenna will receive the reflected signal with nearly as much gain as the direct signal.

3.4.6 Antenna pattern

GPS antennas for positioning applications are typically installed upright (boresight facing zenith). The RHCP gain pattern is very much omni-directional in azimuth and quasi-hemispherical in elevation angle. The LHCP gain pattern is not as well-defined, except that near boresight it is much smaller than RHCP (by ~ 20 dB); in the antenna anti-boresight direction, there are alternating regions where RHCP and LHCP predominate (Figure 12).

Comparing different geodetic antennas (Figure 13), there is no significant difference in the interferometric power (i.e., reflected over direct) for a soil surface. It is only for metallic surfaces that the classic choke-ring design outperforms the other antennas in terms of multipath mitigation (Figure 14). Interestingly, a metallic horizontal surface appears less harmful than bare ground at near-grazing incidence, since geodetic-quality antennas are designed to reject LHCP reflections, but offer little impediment for RHCP intake.

The antenna phase pattern may also impact multipath. At RHCP, this effect is at the ~ mm level for geodetic antennas, and thus usually negligible. The phase difference across different polarizations at the same viewing direction is approximated here as -90°, as discussed in Zavorotny et al. (2010). This has a greater impact for normal incidence or conducting surfaces, which produce predominantly LHCP reflections. It is crucial for replicating the change in SNR

modulation phase observed between dry and wet soil conditions. The LHCP phase pattern needs better characterization in the future.

3.4.7 Antenna orientation

For positioning applications the ideal antenna installation is upright (zenith-looking). However, some reflectometry studies tip the antenna (boresight facing the horizon) or have it upside-down (nadir-looking). The main advantage is the improved reception of reflections. These special orientations require a dedicated installation, in contrast to the upright configuration, which allows the GPS unit to be shared with geodesists, surveyors, and atmospheric scientists.

Changing the antenna orientation also allows neglecting polarization diversity under certain circumstances. For example, with a tipped antenna and low-elevation satellite, the LHCP reflection is suppressed; with an upside-down antenna and a high-elevation satellite, the RHCP component is suppressed. Another simplification offered by tipped installations is that the antenna response can be neglected altogether, as it subjects like-polarized direct and reflected signals to practically the same response (Treuhaft et al. 2001). A tipped antenna installation (Rodriguez-Alvarez et al. 2011a) is also advantageous when using shorter code modulations, such as C/A, because it reduces the cross-channel self-interference (as the maximum gain is applied to the rising or setting satellite being pointed at, rather than to a different high-elevation satellite simultaneously in view). When using the longer codes (e.g., L2C) though, tipping the antenna becomes not only less necessary but also disadvantageous, because it loses satellites visible in azimuths far from boresight, which would offer more frequent retrievals.

Other times a more drastic configuration is chosen, with the goal of measuring only the reflection, e.g. an LHEP antenna (i.e., predominantly LHCP) is installed upside-down (Löfgren et al. 2011) (Figure 15). In this case, the absence of interference fringes is caused by a weak

direct voltage, not a weak reflection. The carrier-phase and pseudorange multipath errors are essentially the respective interferometric quantities. The delay error reaches twice the reflector height (1.5 m) at zenith; phase error exceeds 90° – and in fact wraps around the 19-cm wavelength, as intended.

3.5 CONCLUSIONS

We have presented a forward model for GPS terrestrial multipath – i.e., reflections off land, water, etc. – as experienced by near-surface receivers. It produces GPS carrier-phase, pseudorange, and SNR observables, combining different surface and antenna types, all with due consideration for electromagnetic polarization and coherence. The forward model requires *a priori* information about the parameters affecting the amount of attenuation as well as group and phase delay exhibited by reflections, compared to the direct or line-of-sight signal:

- properties of target surface (geometry and composition),
- measurement system characteristics (code modulation, receiver tracking algorithms, antenna radiation patterns), and
- monitoring setup (height of the antenna above the ground, as well as its orientation).

Starting from the direct and reflected voltages, we have defined and related the interferometric and error voltages. We presented and discussed a number of scenarios, based on which we have drawn conclusions useful for the design and analysis of reflectrometry experiments. For instance, we considered how the antenna orientation – upright, tipped, or upside-down –, involves a number of trade-offs, regarding the neglect of the antenna gain pattern, the minimization of CDMA self-interference, and the maximization of the number of satellites visible. This forward model was also used to understand the multipath signature in GPS positioning applications. For example, we have shown how geodetic GPS antennas offer little

impediment for the intake of near-grazing reflections off natural surfaces (in contrast to, e.g., metal), because of the lack of diversity with respect to the direct signal, i.e., small interferometric delay and Doppler, like sense of polarization, and similar direction of arrival.

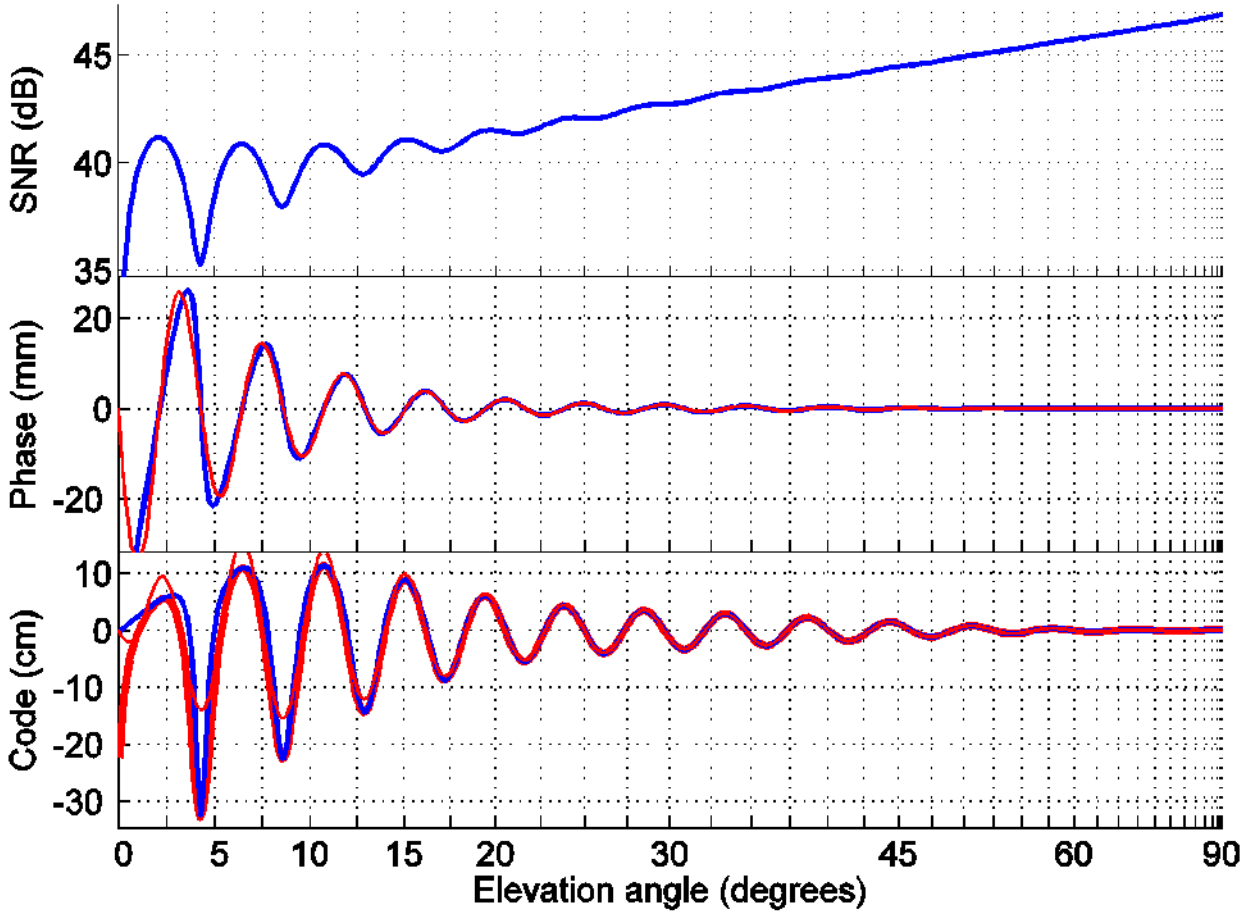

Figure 5: Multipath signature in GPS SNR, carrier phase, and code pseudorange observables for a typical setup. The reflecting surface is horizontal, made up of soil with medium-level moisture and negligible roughness. A choke-ring antenna installed upright on a 1.5-m tall monument is postulated. The rigorous solution is shown in blue, approximations in red; for code pseudorange, the thick (thin) red line corresponds to small-delay (small-delay and small-power) approximation. Please notice the difference in scales between code (cm) and phase (mm)

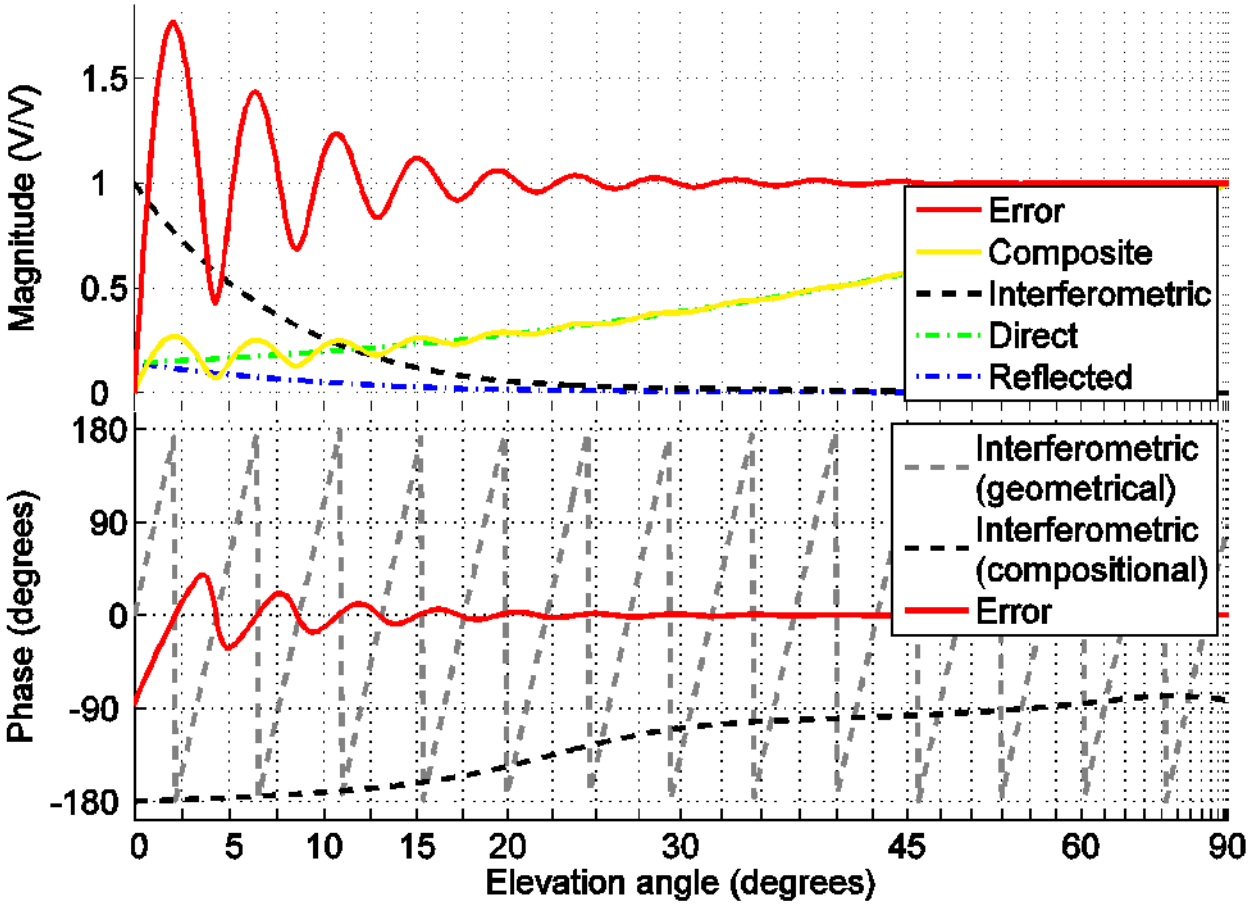

Figure 6: Magnitude and phase of modeled signals. Top panel: reflected, direct, interferometric, composite, and error magnitudes; bottom panel: interferometric and error phases. These are the underlying constituting quantities responsible for producing the observable signatures presented in Figure 5.

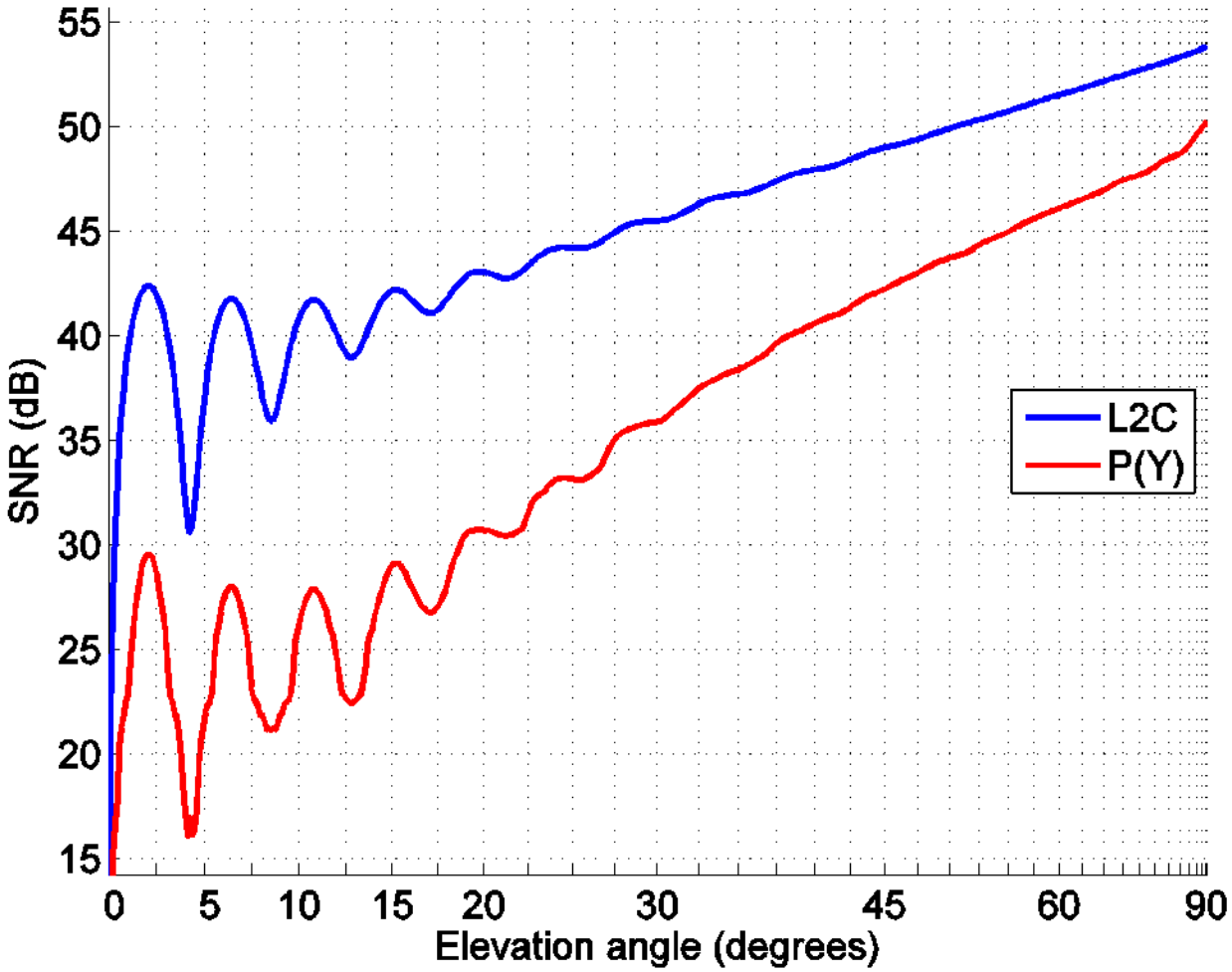

Figure 7: Multipath model results for different code modulations on L2

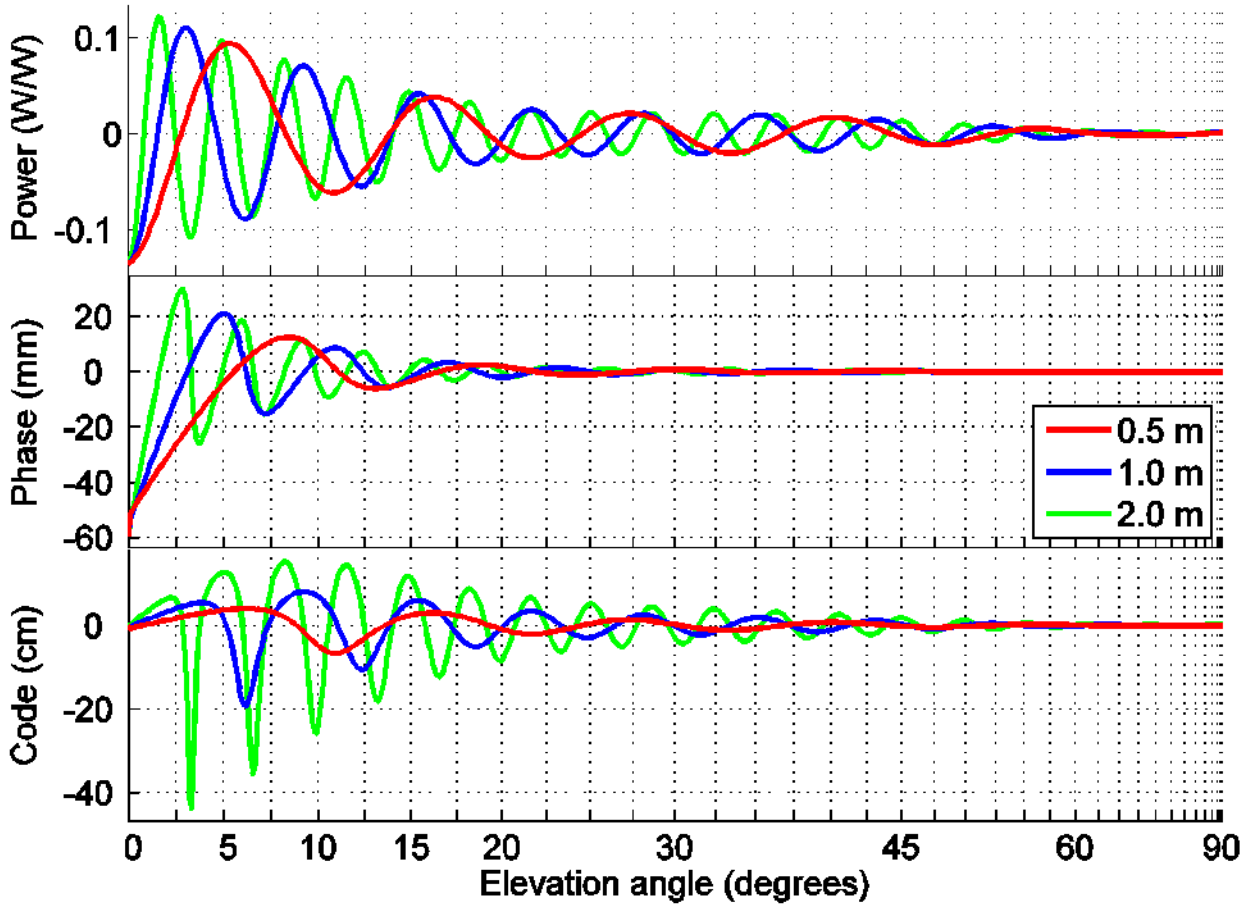

Figure 8: Effect of reflector height on GPS multipath errors. Reflector height values equal to 1/2, 1, and 2 meters are shown in red, blue, and green, respectively

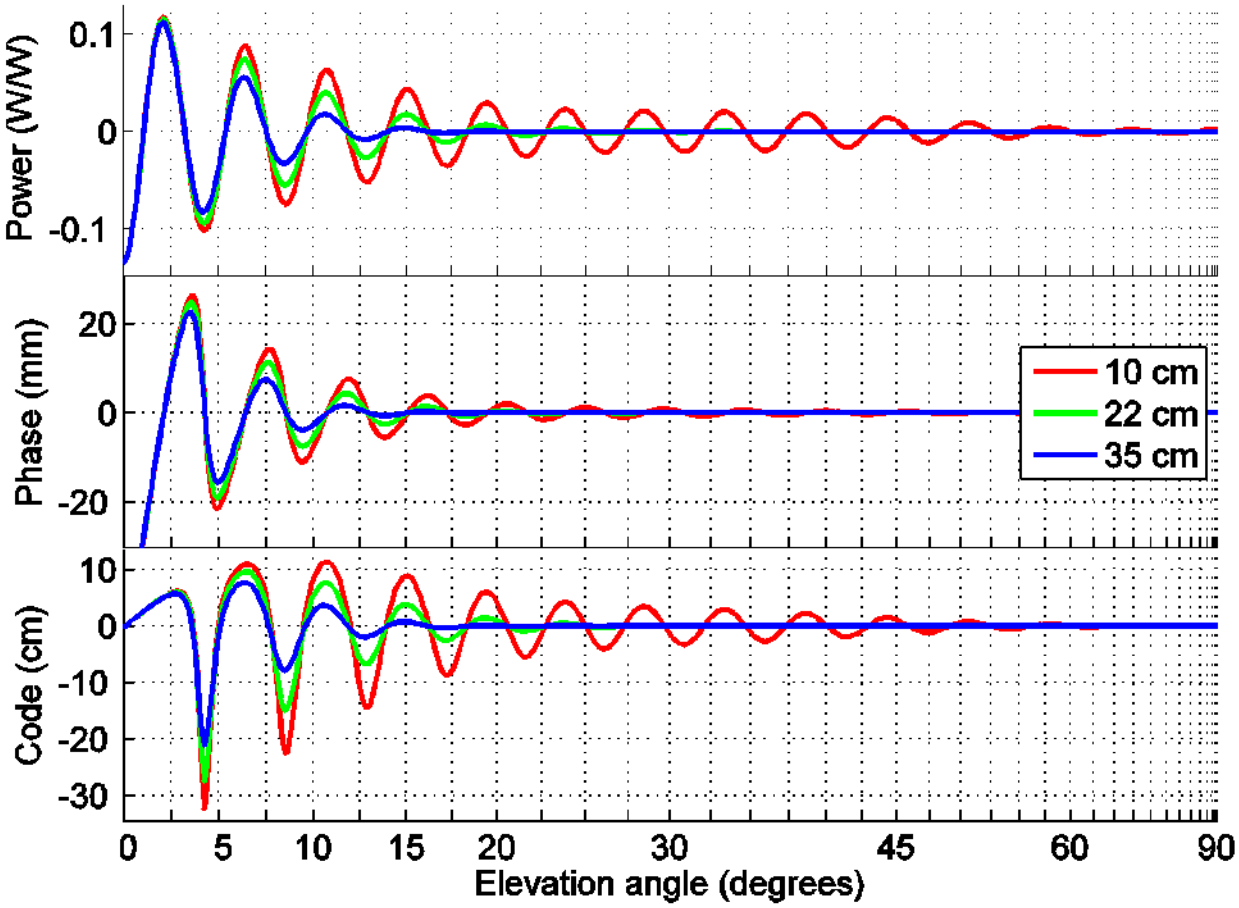

Figure 9: Effect of surface random roughness on GPS multipath errors. Surface height standard deviation values equal to 0, 25, and 35 cm are shown in red, green, and blue, respectively

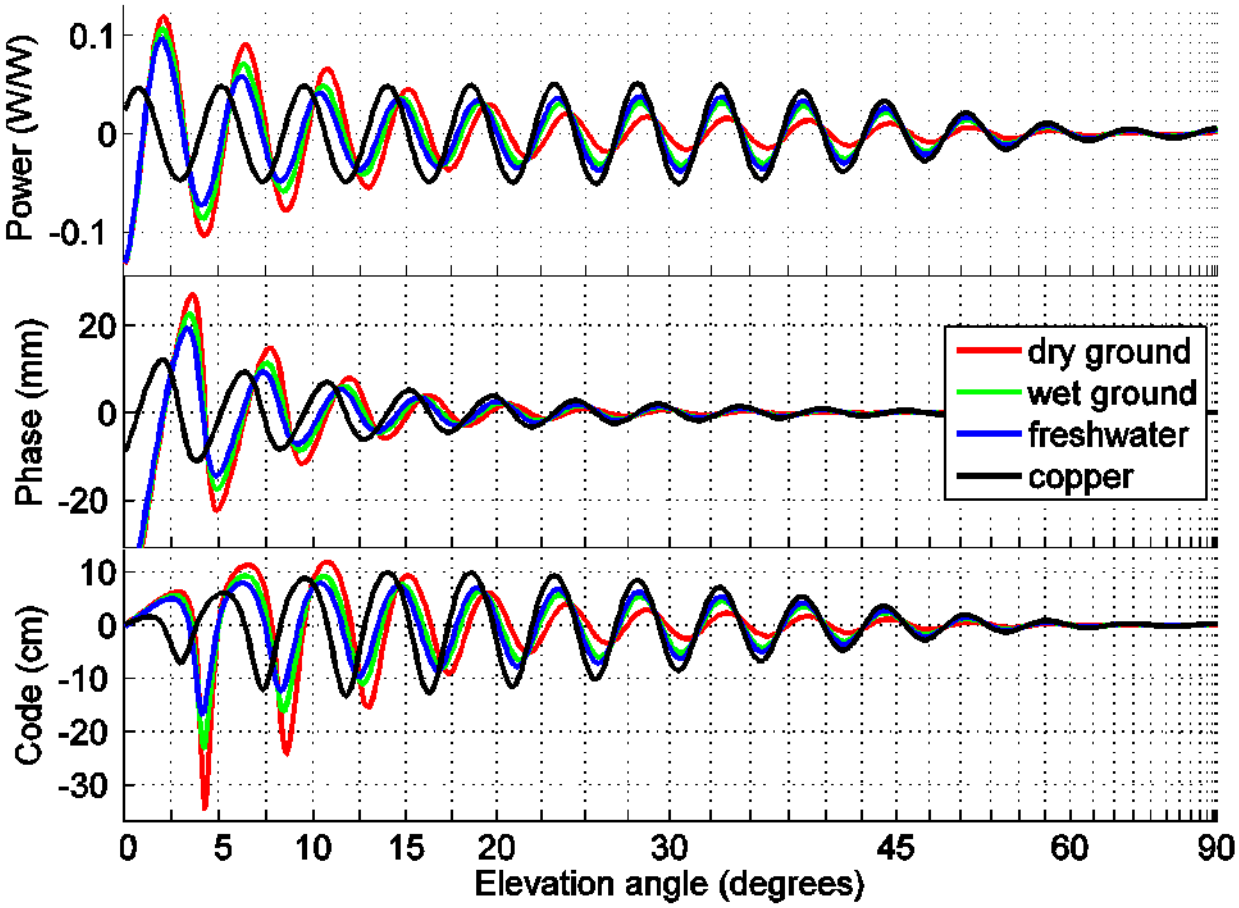

Figure 10: Effect of surface material composition on GPS multipath errors

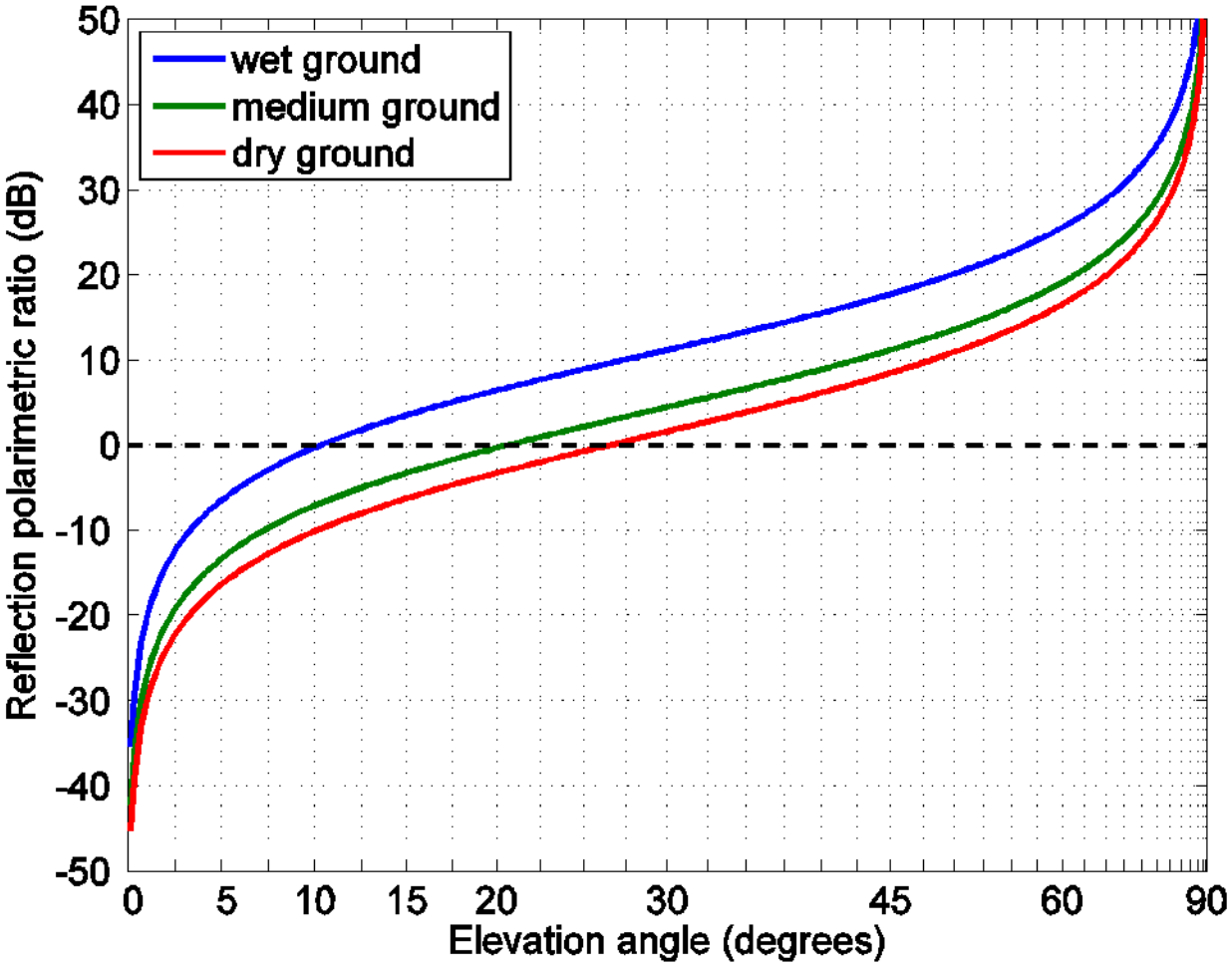

Figure 11: Effect of soil moisture on reflection polarimetric ratio. The Brewster angle is found at the intersection with the horizontal line at zero

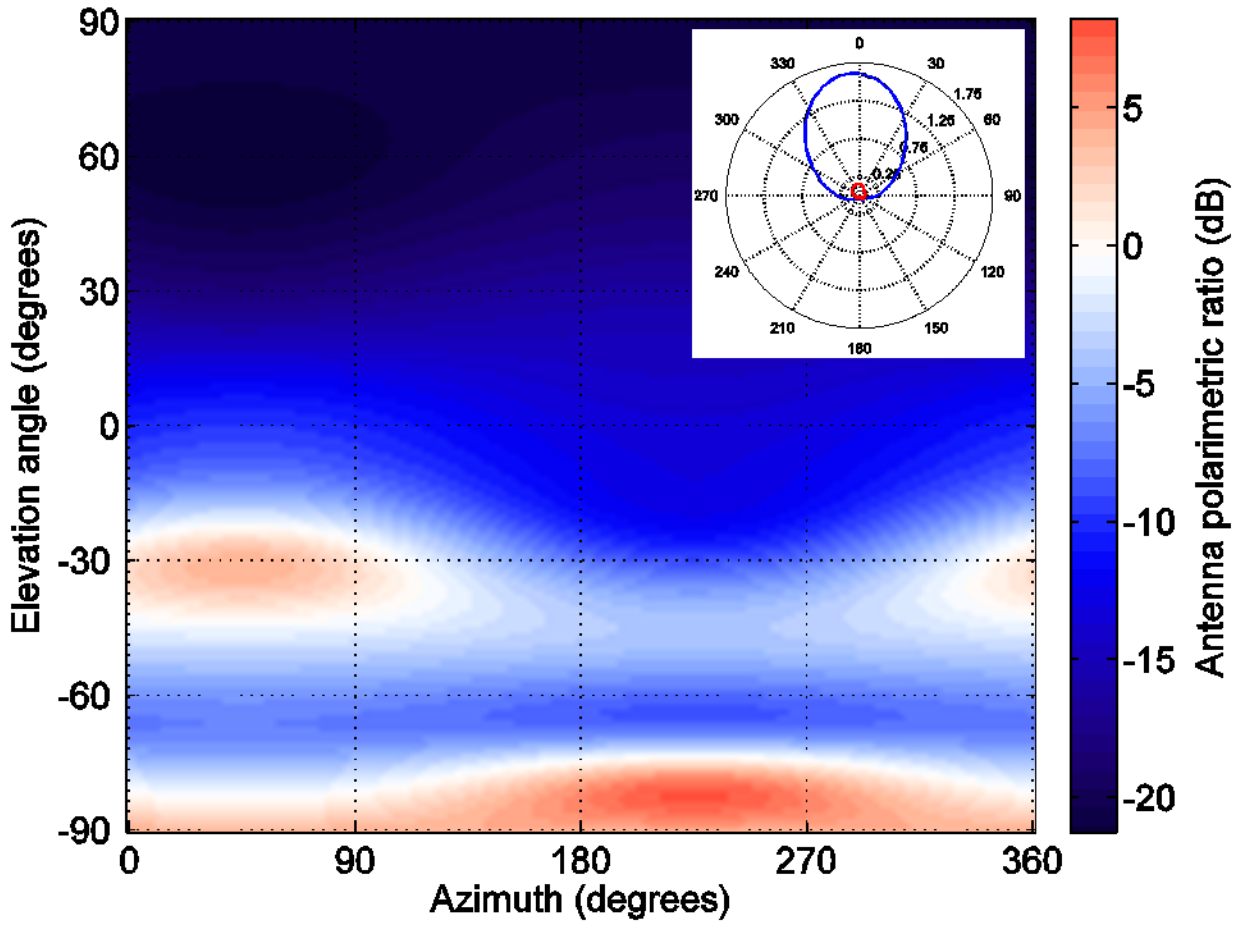

Figure 12: Receiver antenna gain for varying satellite direction. Satellite zenith angle and azimuth coincide with antenna boresight angle and axial angle only in a zenith- and north-aligned setup. Antenna polarimetric ratio is in decibels; negative (positive) values correspond to RHEP (LHEP) and are shown in blue (red). Inset: antenna gain separately at each polarization (units: meters); RHCP is shown in blue, LHCP in red. The two-dimensional grid shown in the main plot is based on a spherical harmonics expansion of the one-dimensional principal plane cut shown in the inset

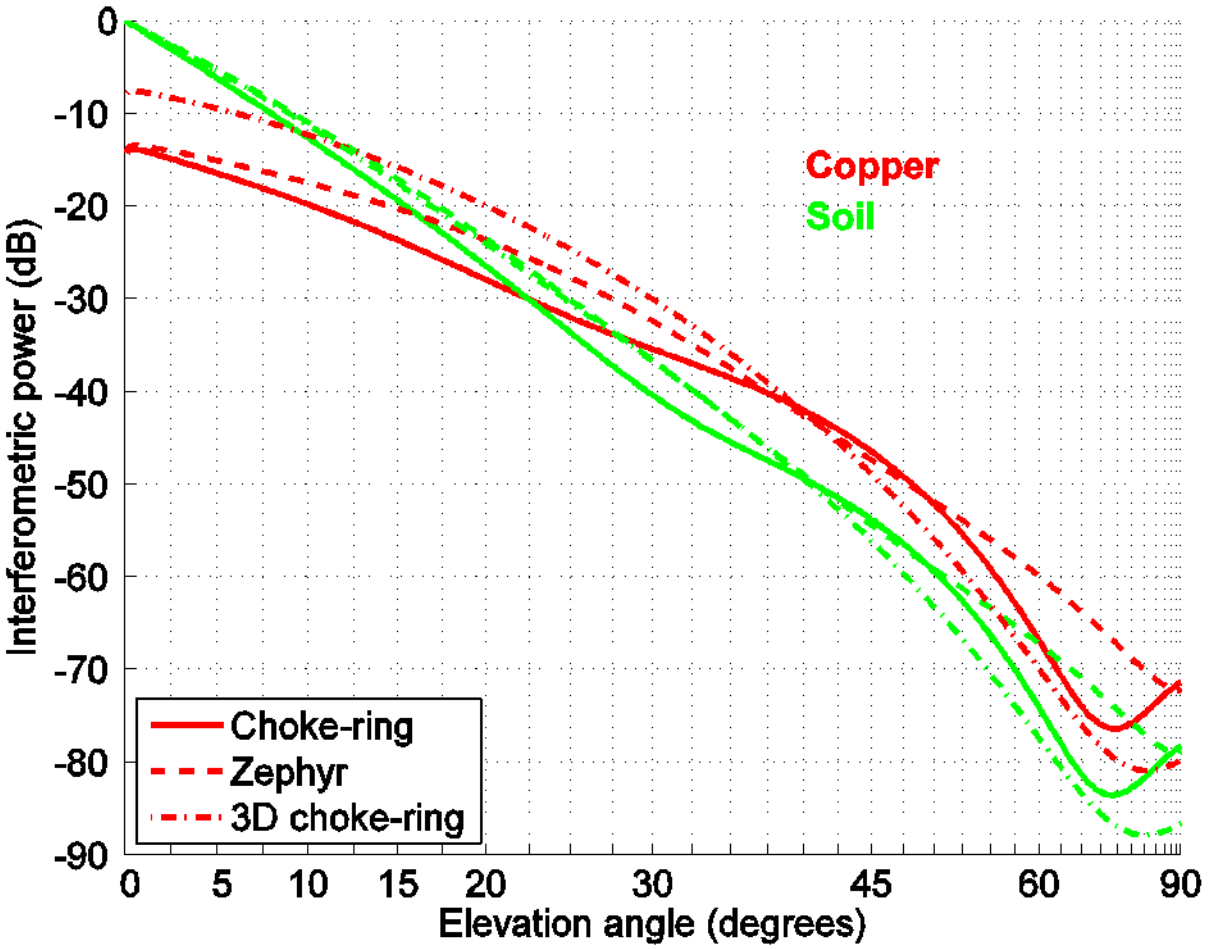

Figure 13: Effect of antenna model on interferometric power. Combinations of two surface materials and three geodetic-quality antenna models are compared. Soil is shown in green, copper in red; choke-ring, zephyr, and 3D choke-ring (IGS antenna codes TRM29659.00, TRM55971.00, and LEIAR25) are shown respectively in solid, dashed, and dash-dot line styles. L1 and C/A are assumed for the carrier frequency and code modulation

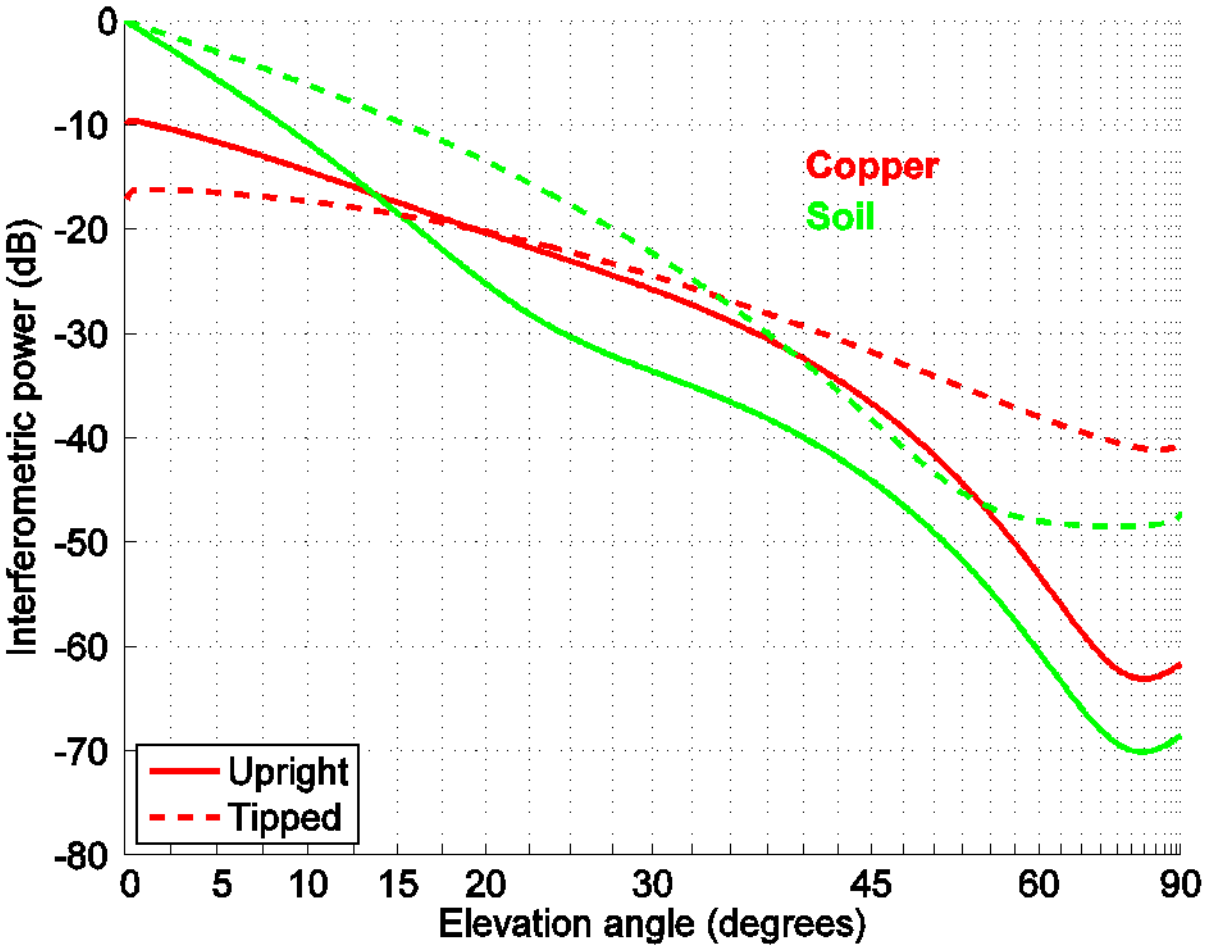

Figure 14: Effect of antenna orientation on interferometric power. Combinations of two surface materials and two antenna orientations are compared. Soil is shown in green, copper in red; upright and tipped orientations are shown respectively in solid and dashed line styles

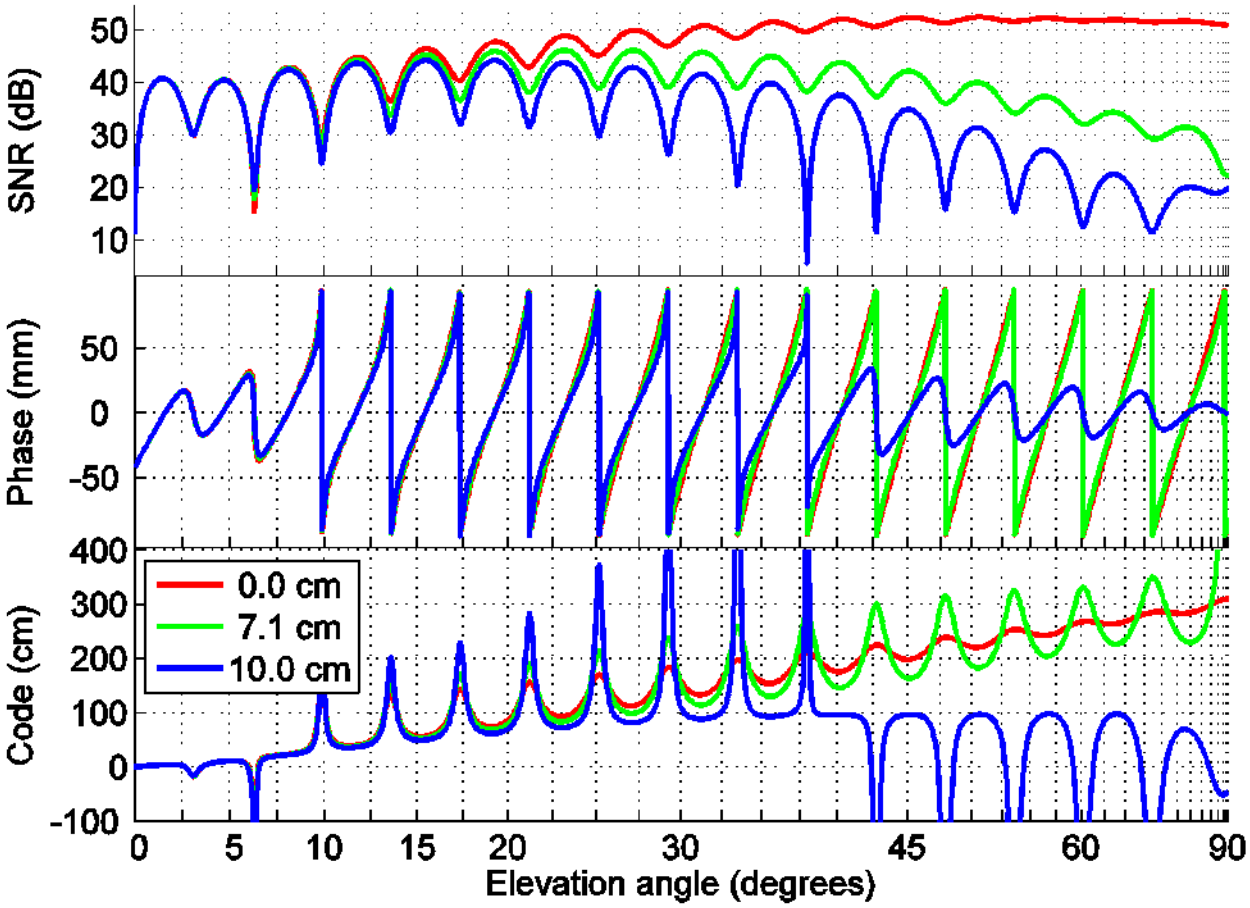

Figure 15: Multipath signature in GPS SNR, carrier phase, and code pseudorange observables for an atypical setup. An LHCP-predominant antenna is installed upside-down 1.5 m above seawater. Results for varying surface random roughness are shown in red, green, and blue, corresponding to surface height standard deviation values of 0, 7.1, and 10.0 cm, respectively. L1 and C/A are assumed for the carrier frequency and code modulation

Chapter 4: Inverse modeling of GPS multipath for snow depth estimation – Part I: Formulation and simulations

4.1 INTRODUCTION

This is part I of a contribution that uses a statistical inverse model for GPS multipath to estimate snow depth. We apply the method of non-linear least squares to retrieve parameter corrections – both their most probable values as well as their uncertainty – from the measurement/simulation residuals. We have developed a parameterization with sufficient statistical degree of freedom so as to mitigate noise while seeking minimum-variance and unbiased parameter estimates. Given adequate external constraints, the information content of GPS measurements is reserved for the determination of the environmental parameters of interest, such as snow depth, instead of nuisance parameters such as antenna gain patterns.

After briefly reviewing the forward model, we proceed to explain how the unknown snow characteristics were parameterized and how these biases were embedded in the physical forward model. We also explain how we obtained starting values for the parameters, as required for initializing the non-linear optimization. In section 4 we detail the inverse model per se, illustrating the observation/parameter sensitivity and how these are modified by measurement errors. Section 5 we assess the inversion performance employing simulations, in terms of actual errors (retrieval minus true parameters) and to what extent these are bounded by the parameters formal uncertainty. We close the chapter quantifying the dependence of model results on the satellite elevation angle.

4.2 PHYSICAL FORWARD MODEL REVIEW

We briefly summarize the forward model. SNR observations are formulated as:

$$SNR = P_{\mathrm{c}}/P_{\mathrm{n}}. \tag{44}$$

In the denominator we have the noise power P_{n}, here taken as a constant, based on nominal values for the noise power spectral density and the noise bandwidth. The numerator is composite power:

$$P_{\mathrm{c}} = P_{\mathrm{c}}^{\mathrm{I}} + P_{\mathrm{c}}^{\mathrm{C}} \tag{45}$$

Its incoherent component is simply the sum of the respective direct and reflected powers:

$$P_{\mathrm{c}}^{\mathrm{I}} = P_{\mathrm{d}}^{\mathrm{I}} + P_{\mathrm{r}}^{\mathrm{I}} \tag{46}$$

while the coherent component follows from the complex sum of direct and reflection average voltages, $P_{\mathrm{c}}^{\mathrm{C}} = |\langle V_{\mathrm{d}}\rangle + \langle V_{\mathrm{r}}\rangle|^2$:

$$P_{\mathrm{c}}^{\mathrm{C}} = P_{\mathrm{d}}^{\mathrm{C}} + P_{\mathrm{r}}^{\mathrm{C}} + 2\sqrt{P_{\mathrm{d}}^{\mathrm{C}}}\sqrt{P_{\mathrm{r}}^{\mathrm{C}}}\cos\phi_{\mathrm{i}}, \tag{47}$$

expressed in terms of the coherent powers $P_{\mathrm{d}}^{\mathrm{C}} = |\langle V_{\mathrm{d}}\rangle|^2$ and $P_{\mathrm{r}}^{\mathrm{C}} = |\langle V_{\mathrm{r}}\rangle|^2$, as well as the interferometric phase:

$$\phi_{\mathrm{i}} = \arg(\langle V_{\mathrm{r}}\rangle/\langle V_{\mathrm{d}}\rangle) \ = \phi_{\mathrm{r}} - \phi_{\mathrm{d}}, \tag{(48}$$

which amounts to the reflection excess phase with respect to the direct signal. We decompose observations $\mathrm{SNR} = \mathrm{tSNR} + \mathrm{dSNR}$ into a trend

$$\mathrm{tSNR} = (P_{\mathrm{d}} + P_{\mathrm{r}})P_{\mathrm{n}}^{-1} = P_{\mathrm{c}}^{\mathrm{I}}P_{\mathrm{n}}^{-1} + \left(P_{\mathrm{d}}^{\mathrm{C}} + P_{\mathrm{r}}^{\mathrm{C}}\right)P_{\mathrm{n}}^{-1} \tag{49}$$

over which interference fringes are superimposed:

$$\mathrm{dSNR} = 2\sqrt{P_{\mathrm{d}}^{\mathrm{C}}}\sqrt{P_{\mathrm{r}}^{\mathrm{C}}}P_{\mathrm{n}}^{-1}\cos\phi_{\mathrm{i}}. \tag{50}$$

From now on we neglect incoherent composite power $P_{\mathrm{c}}^{\mathrm{I}}$ – which only impacts tSNR, not dSNR –, and drop the superscript for coherent powers.

The direct or line-of-sight power is formulated as

$$P_{\mathrm{d}} = P_{\mathrm{d}}^{\mathrm{R}}G_{\mathrm{d}}^{\mathrm{R}}, \tag{51}$$

where $P_{\mathrm{d}}^{\mathrm{R}}$ is the direction-dependent right-hand circularly polarized (RHCP) power component incident on an isotropic antenna, as specified in the GPS interface [GPSD-USAF, 2011]; the

same document specifies the left-handed circularly polarized (LHCP) component to be negligible. The direct antenna gain $G_{\mathrm{d}}^{\mathrm{R}}$ is obtained evaluating the antenna pattern at the satellite direction and at RHCP polarization. For most commonly used antennas, the full pattern is made available by the antenna manufacturing company.

The reflection power,

$$P_{\mathrm{r}} = P_{\mathrm{d}}^{\mathrm{R}} |X|^2 S^2, \tag{52}$$

is defined starting with the same incident power $P_{\mathrm{d}}^{\mathrm{R}}$ as in the direct power P_{d}. It ends with a coherent power attenuation factor caused by random surface roughness,

$$S^2 = \exp(-k^2 s^2 \cos^2 \theta), \tag{53}$$

where θ is the angle of incidence (with respect to the surface normal), $k = 2\pi/\lambda$ is the wavenumber, and $\lambda = 24.4$ cm is carrier wavelength for L2C. The effective roughness, denoted s (in meters), represents the surface height standard deviation with respect to the spatially non-uniform surface trend. At the core of P_{r} we have coupled surface/antenna reflection coefficients, $X = X^{\mathrm{R}} + X^{\mathrm{L}}$, defined here as

$$X^{\mathrm{R}} = R^{\mathrm{S}} \sqrt{G_{\mathrm{r}}^{\mathrm{R}}} \exp(\iota \Phi_{\mathrm{r}}^{\mathrm{R}}), \tag{54}$$

$$X^{\mathrm{L}} = R^{\mathrm{X}} \sqrt{G_{\mathrm{r}}^{\mathrm{L}}} \exp(\iota \Phi_{\mathrm{r}}^{\mathrm{L}}), \tag{55}$$

producing respectively RHCP and LHCP fields (with the imaginary unity denoted $\iota = \sqrt{-1}$). The antenna response includes the power gain G and the phase center variation Φ, evaluated at the reflection direction, and separately for each polarization. The surface response is represented by the complex-valued Fresnel coefficients, $R^{\mathrm{X}}, R^{\mathrm{S}}$, for cross- and same-sense circular polarization, respectively. The medium is assumed to be homogeneous (i.e., a semi-infinite half-space). We allow for varying material types (e.g., snow and soil) and material properties (snow

density and soil moisture moisture). Material models provide the complex permittivity which drives the Fresnel coefficients.

We neglect the following effects: random volumetric inhomogeneities as well as internal layers – i.e., both discrete interfaces (such as snow/ground) and continuously-varying stratification (e.g., from lower to higher density); changes in reflection power due to ray focusing/spreading on a concave/convex surface; phase changes caused by coordinate basis differences (between surface- and antenna-aligned bases); and the CDMA modulation impressed on the carrier wave, which is acceptable for small interferometric delay and Doppler, such as in the case of grazing incidence, stationary surface/receiver conditions, and short antenna installations.

The interferometric phase reads:

$$\phi_{\mathrm{i}} = \phi_X + \phi_I - \Phi_{\mathrm{d}}^{\mathrm{R}}. \tag{56}$$

The first term $\phi_X = \arg(X)$ accounts for the surface and antenna properties, as above. The last one $\Phi_{\mathrm{d}}^{\mathrm{R}}$ is the direct phase contribution, which amounts to only the RHCP antenna phase center variation evaluated at the satellite direction. The majority of the components present in the direct RHCP phase ϕ_{d} (such as receiver and satellite clock states, the bulk of atmospheric propagation delays, etc.) are also present in the reflection phase ϕ_{r}, so they cancel out in forming $\phi_{\mathrm{i}} = \phi_{\mathrm{r}} - \phi_{\mathrm{d}}$, which is a very fortunate simplification.

At the core of the interferometric phase ϕ_{i} we have the geometric component $\phi_I = k\tau_{\mathrm{i}}$, the product of the wavenumber k (in radians per meter) and the interferometric propagation delay τ_{i} (in meters). Assuming a locally horizontal surface, the latter is simply (Georgiadou and Kleusberg 1988):

$$\tau_{\mathrm{i}} = 2H_A \sin e, \tag{57}$$

in terms of the satellite elevation angle e and an *a priori* reflector height H_A

4.3 PARAMETERIZATION OF UNKNOWNS

There are errors in the nominal values assumed for the physical parameters. Ideally we would estimate separate corrections for each one. Unfortunately many are linearly dependent or nearly so. Because of this dependency, we have kept physical parameters fixed to their optimal *a priori* values, and estimated a few biases (detailed below). Each is an amalgamation of corrections for different physical effects. In a later stage, we rely on multiple independent bias estimates (e.g., successive days) to try and separate the physical sources.

4.3.1 Biases

The original noise power P_{n} will be augmented as P_{n}/K to include a bias K accounting for imperfections in the nominal trend of direct power vs. elevation angle, $P_{\mathrm{d}}^{\mathrm{R}}$ – including unanticipated attenuations along the line-of-sight, such as foliage, precipitation, etc. –, and also in the nominal values of all other linearly-dependent parameters, mainly antenna gain $G_{\mathrm{d}}^{\mathrm{R}}$ (direction-dependent) and the nominal noise power P_{n} (direction-independent; it is temperature-dependent).

The noise power bias K must be nonnegative, yet the optimization is performed over the real numbers, so we handle this bias expressed in decibels, $K_{\mathrm{dB}} = 10 \log_{10} K$. Furthermore, we expand it as a polynomial in terms of powers of sine of elevation angle $\sin e$:

$$K_{\mathrm{dB}} = K_{\mathrm{dB}}^{(0)} + K_{\mathrm{dB}}^{(1)} \sin e + K_{\mathrm{dB}}^{(2)} \sin^2 e + \cdots = \sum_{j=0,1,\ldots} K_{\mathrm{dB}}^{(j)} \sin^j e. \tag{58}$$

We also introduce an elevation angle dependent bias function $B = |B| \exp(\iota \phi_B)$, as a correction for imperfections in the reflection model. Reflection power is updated as $P_{\mathrm{r}}/|B|^2$ and interferometric phase as $\phi_{\mathrm{i}} - \phi_B$. This complex-valued reflection bias B will compensate for

both errors of commission (e.g., surface roughness that, although accounted for, is over- or under-estimated) and errors of omission (e.g., volumetric inhomogeneities that are completely ignored). Similar as for the noise power bias, we define polynomials in $\sin e$ for the reflection power bias $B_{\mathrm{dB}} = \sum_{j=0,1,\ldots} B_{\mathrm{dB}}^{(j)} \sin^j e$ and for the reflection phase bias $\phi_B = \sum_{j=0,1,\ldots} \phi_B^{(j)} \sin^j e$.

We single out the first two phase bias coefficients. The constant term $\phi_B^{(0)} = \varphi_B$ is a phase-shift accounting for errors in the antenna phase pattern Φ, in the medium composition, surface tilting, etc. The linear phase coefficient gets recast as an equivalent horizontal-surface reflector height,

$$H_B = (4\pi)^{-1} \lambda \phi_B^{(1)}. \tag{59}$$

It accounts for errors in the *a priori* value for the reflector height, including the unknown thickness of a snow layer deposited over the ground and a tilting in the underlying ground (yielding a positive topographical height bias downhill and negative uphill). The full phase bias function is rewritten as

$$\phi_B = \varphi_B + 4\pi\lambda^{-1} H_B \sin e + \hat{\phi_B} \tag{60}$$

where the phase remainder $\hat{\phi_B} = \sum_{j=2,3,\ldots} \phi_B^{(j)} \sin^j e$ contains higher-order terms, $j = 2,3$.

The forward model, including biases, can be summarized as follows:

$$\mathrm{SNR} = \left(1 + P_{\mathrm{i}} + 2\sqrt{P_{\mathrm{i}}} \cos \phi_{\mathrm{i}}\right) P_{\mathrm{d}} K / P_{\mathrm{n}} + P_{\mathrm{c}}^{\mathrm{I}} / P_{\mathrm{n}} \tag{61}$$

where interferometric power and phase are, respectively:

$$P_{\mathrm{i}} = |XS/B|^2 / G_{\mathrm{d}}^{\mathrm{R}}, \tag{62}$$

$$\phi_{\mathrm{i}} = \phi_X + 4\pi\lambda^{-1} H \sin e - \varphi_B - \hat{\phi_B} - \Phi_{\mathrm{d}}^{\mathrm{R}}. \tag{63}$$

The total reflector height $H = H_A - H_B$ (*a priori* value minus unknown bias) is to be interpreted as an effective value that best fits measurements, which includes snow and other components.

4.3.2 Observation partitioning and parameters update rate

SNR observations are recorded at regular time intervals (originally 1-second, decimated to 10 seconds) for different satellites tracked from the same station. Satellite azimuth and elevation angle are calculated from the broadcast orbits and serve as independent variables guiding the partitioning of observations into tracks. Elevation angle is the one variable over which the model varies the most: reflection coefficients, surface roughness, interferometric propagation delay, antenna gain pattern. Therefore, it is desirable to maximize the range of observed elevation angles per track, if the various parameters being jointly estimated are to be resolved without ambiguity. Azimuth, on the other hand, offers little for such an ambiguity resolution (assuming a planar horizontal surface); therefore, we wish to minimize the azimuth range per track, to avoid the risk of conflating distinct scattering conditions (e.g., presence vs. absence of trees, varying snow depth due to different orientation to the Sun, etc.).

We partition the data into ascending and descending tracks, between a satellite's rise and set on the horizon, split at the satellite culmination, near zenith. Each track has a duration of ~ 1-2 hours. This configuration normally offers a sufficient range of elevation angles, unless the satellite reaches culmination too low in the sky ($\lesssim 20°$), in which case that arc is discarded. In seeking a balance between under- and over-fitting – between an insufficient and an excessive number of parameters –, we found it appropriate to estimate the following unknown parameters:

$$\mathrm{X} = \left[H_B, \varphi_B, B_{\mathrm{dB}}^{(0)}, B_{\mathrm{dB}}^{(1)}, B_{\mathrm{dB}}^{(2)}, K_{\mathrm{dB}}^{(0)}, K_{\mathrm{dB}}^{(1)}, K_{\mathrm{dB}}^{(2)}\right]^{\mathrm{T}}. \tag{64}$$

Figure 16 shows the effect of the constant and linear biases on the SNR observations. Reflector height bias H_B changes the number of oscillations; phase-shift φ_B displaces the oscillations along the horizontal axis; reflection power $B_{\mathrm{dB}}^{(0)}$ affects the depth of fades; zeroth-order noise power $K_{\mathrm{dB}}^{(0)}$ shifts the observations up or down as a whole; and first-order noise power $K_{\mathrm{dB}}^{(1)}$ tilts

the SNR curve. A good parameterization yields observation sensitivity curves as unique as possible for each parameter.

4.3.3 Bootstrapping parameter priors

Biases and SNR observations are involved non-linearly through the forward model. Therefore, there is the need for a preliminary global optimization, without which the subsequent final local optimization will not necessarily converge to the optimal solution. Its possibly coarse estimates will be refined once the inversion is bootstrapped.

Figure 17 shows slices of the multi-dimensional parameteric space over which the optimization takes place. Such exhaustive gridded sampling of the minimum of residuals is done only for illustration purposes and is unnecessary in an efficient inversion. The reflector height/phase shift subspace has a clear minimum at the true solution; the other minima at $\pm 360°$ are mirrored copies of the main one, which are innocuous as the phase domain is cyclic. The reflector height/reflection power subspace also exhibits a basin of attraction guaranteeing convergence in the immediate vicinity to the true solution. In fact, provided the initial height is correct, the initial reflection power bias is not very important, as residuals form a simple corner along the latter dimension. Yet away from the true reflector height, reflection power bias loses its well-defined minimum; instead, residuals slope downward towards decreasing power. Intuitively, given initial fringes of the wrong height, the optimization will prefer to fit no fringe at all. Finally, in the reflector height/noise power subspace, there is a similar basin surrounded by ridges beyond which convergence is not guaranteed. A reasonable initial height again makes the initial noise power mostly irrelevant, because of the corner in residuals along the latter dimension.

SNR observations would trace out a perfect sinusoid curve in the case of an antenna with isotropic gain and spherical phase pattern, surrounded by a smooth, horizontal, and infinite surface (free of small-scale roughness, large-scale undulations, and edges), made of perfectly electrically conducting material, and illuminated by constant incident power. Thus in such an idealized case, taking $4\pi\lambda^{-1}\sin e$ as the independent variable, SNR could be described exactly by constant reflector height, phase-shift, amplitude, and mean values. As the measurement conditions become more complicated, SNR starts to deviate from a pure sinusoid.

A polynomial/spectral decomposition often remains adequate. We illustrate the procedure using a representative measurement sample, as follows. First we simulate a trend, tSNR', free of interference fringes, by artificially forcing $\phi_{\mathrm{i}} = 90°$ so that $\cos\phi_{\mathrm{i}}$ is nullified. The ratio between measured and simulated observations, $\mathrm{rSNR} = \mathrm{SNR}/\mathrm{tSNR}'$, provides trend residuals (Figure 18). These are fit by a low-order polynomial. The resulting noise power bias $K = 10^{K_{\mathrm{dB}}/10}$ is subsequently employed to update the trend simulation, tSNR'. Second, we detrend measurements, $\mathrm{dSNR} = \mathrm{SNR} - \mathrm{tSNR}'$, as well as simulations, $\mathrm{dSNR}' = \mathrm{SNR}' - \mathrm{tSNR}'$, which leaves just the fringes (Figure 19). We then fit a sinusoid separately to each dSNR and dSNR' (see below for details). This yields two sets of reflector height, phase-shift, and amplitude: H, φ, A for measurements; and H', φ', A' for simulations. Finally, we subtract the simulated results from the measured ones:

$$H_B = H' - H, \tag{65}$$

$$\varphi_B = \varphi' - \varphi, \tag{66}$$

$$|B| = A/A'. \tag{67}$$

Clearly, the normalization $|B| = A/A'$ is crucial if reflection power estimates are to be independent of the particular antenna gain utilized. By applying the same spectral analysis to

measured and simulated SNR and then differencing the two sets we cancel out the bulk of both, resulting in approximate values for the unknown biases (H_B, φ_B, $|B|$). These values are used as initial or prior values for the zeroth-order polynomial coefficients, e.g., $B_{\mathrm{dB}}^{(0)} = \log_{10}(20\,|B|)$; higher-order parameters, if present, are initialized as zero.

The best-fitting sinusoid is found by seeking the peak amplitude in a Lomb-Scargle periodogram, also known as least-squares spectral analysis. It is preferable to a more commonly used FFT, given irregularly spaced measurements. It requires a pre-determined domain of trial reflector heights, encompassing the full range of expected conditions (including snow and topography), spaced at the desired precision.

4.4 STATISTICAL INVERSE MODEL FORMULATION

Based on the preliminary values for the unknown parameters vector X and other known (or assumed) values, we run the forward model to obtain simulated observations, $f(\mathrm{X})$. By comparing $f(\mathrm{X})$ against measurements Y we form pre-fit residuals, $\mathrm{y} = \mathrm{Y} - f(\mathrm{X})$. These in turn serve to retrieve parameter corrections,

$$\hat{\mathrm{x}} = \mathrm{C}_{\hat{\mathrm{x}}}\mathrm{J}^{\mathrm{T}}\mathrm{C}_{\mathrm{y}}^{-1}\mathrm{y}, \tag{68}$$

such that the squared sum $\hat{\mathrm{y}}^{\mathrm{T}}\hat{\mathrm{y}}$ of post-fit residuals $\hat{\mathrm{y}} = \mathrm{Y} - f(\mathrm{X} - \hat{\mathrm{x}})$ is minimized. The impact of matrices $\mathrm{J}, \mathrm{C}_{\mathrm{y}}, \mathrm{C}_{\hat{\mathrm{x}}}$ is discussed below.

4.4.1 Functional model

The Jacobian matrix J is the main component in the inversion. It represents the sensitivity of observations to parameter changes:

$$\mathrm{J}(i,j) = \partial \mathrm{Y}_i / \partial \mathrm{X}_j \tag{69}$$

where the partial derivative is defined element-wise. Instead of deriving analytical expressions, we evaluate them numerically, via finite differencing.

Figure 20 shows each column of the Jacobian matrix as a line trace. The sensitivities with respect to reflector height and phase shift are aligned with each other and exhibit respectively an increasing and decreasing magnitude over increasing elevation angle. In fact, were SNR observations a perfect sinusoid, the two sensitivities would be mirrored copies of each other, coinciding at the central elevation angle. The longer the track, the more distinct the two sensitivity curves are. Conversely, the shorter the track, the more difficult it is to distinguish reflector height and phase shift effects. Consequently, their parameters estimates will exhibit significant correlation which in turn will deteriorate the reflector height uncertainty. In fact, were it possible to account for phase shift in the forward model and exclude or at least tightly constrain it in the inverse model, reflector height precision would be greatly improved. Conversely, constraining reflector height would improve the precision of phase shift estimates.

The sensitivity with respect to reflected power bias – essentially a scaled version of the dentrended SNR observations – is in phase quadrature with that of reflector height and phase shift. This orthogonality implies that the former can be decorrelated well from the latter two parameters. This is not to say that reflection power has no impact on the reflector height uncertainty. The sensitivity of observations with respect to reflector height depends not only on reflector height itself but potentially also on all other parameters. Indeed, if the reflection power becomes too small, simulated observations lose sensitivity to reflector height: $\lim_{B_{\mathrm{dB}}\to-\infty} \partial \mathrm{Y}/\partial H_B = 0$. This may be caused by a genuine physical effect, as in the case of random surface roughness; other times, it is an undesirable numerical artifact arising from an inadequate optimization (see below for ways to mitigate this via the stochastic model).

Finally, the sensitivities with respect to noise power would in principle be simply polynomial bases, i.e., zeroth and first-order powers of sine of elevation angle. Yet these

sensitivities fail to exhibit a purely constant or linear behavior. That is because the polynomial is defined in decibels (necessary to guarantee non-negative power). In contrast, we take observations in watts per watt, because these are the units in which fringes are supposed to most resemble a sinusoid. The estimation of noise power bias only affects the reflector height uncertainty for very small total reflector heights; in this case both sensitivities will exhibit large-period oscillations, resembling a low-order polynomial. Thus the impact arises from the near co-linearity between the two parameters, not from the noise power sensitivity alone.

4.4.2 Stochastic model

We need to specify the uncertainty and correlation expected in residuals and parameters. Their prior covariance matrices modify the objective function being minimized, which now becomes

$$\mathrm{SSR} = \hat{\mathrm{x}}^{\mathrm{T}}\mathrm{C}_{\mathrm{x}}^{-1}\hat{\mathrm{x}} + \hat{\mathrm{y}}^{\mathrm{T}}\mathrm{C}_{\mathrm{y}}^{-1}\hat{\mathrm{y}} \tag{70}$$

(instead of just $\hat{\mathrm{y}}^{\mathrm{T}}\hat{\mathrm{y}}$); the least-squares solution becomes a fusion of prior information and new measurements.

The residuals prior covariance matrix C_{y} cannot be simply a multiple of the identity matrix. SNR measurements are reported in decibels, which yields a logarithm scaling in watts per watt. Starting with a homogeneous residual variance in decibels $\sigma^2_{\mathrm{y}_{i\mathrm{dB}}}$, we obtain heterogeneous variances via the delta method:

$$\sigma^2_{\mathrm{y}_i} = \sigma^2_{\mathrm{y}_{i\mathrm{dB}}}(\partial \mathrm{Y}_i/\partial \mathrm{Y}_{i\mathrm{dB}})^2, \tag{71}$$

So variances are scaled in proportion to the decibel-to-power transformation itself,

$$\partial \mathrm{Y}_i/\partial \mathrm{Y}_{i\mathrm{dB}} = \mathrm{SNR}_i \ln(10)\,/10 \tag{72}$$

where $\mathrm{SNR}_i = 10^{\mathrm{SNR}_{i\mathrm{dB}}/10}$. We form a diagonal residual covariance matrix $\mathrm{C}_{\mathrm{y}} = \mathrm{Diag}\left(\Sigma_{\mathrm{y}}\right)$ given the vector $\Sigma_{\mathrm{y}} = \left[\sigma^2_{\mathrm{y}_1}, \sigma^2_{\mathrm{y}_2}, \ldots\right]^{\mathrm{T}}$ of transformed variances over each i-th satellite direction.

Such a data transformation is used to make the observations conform to the assumptions of the statistical method.

A non-identity prior residual covariance matrix drastically alters the observation/ parameter sensitivities. The impact can be demonstrated rewriting the right-hand side of the normal equations, $\mathrm{J}^{\mathrm{T}}\mathrm{C}_{\mathrm{y}}^{-1}\mathrm{y} = \mathrm{J}'^{\mathrm{T}}\mathrm{y}'$ in terms of a modified Jacobian $\mathrm{J}' = \mathrm{C}_{\mathrm{y}}^{-1/2}\mathrm{J}$ and a modified residual vector $\mathrm{y}' = \mathrm{C}_{\mathrm{y}}^{-1/2}\mathrm{y}$. The pre-multiplication by the matrix square root implies that each original row is scaled by a weight of the form $\sigma_{\mathrm{y}_i}^{-1}$. Columns of the modified Jacobian are shown in Figure 21. The noise power parameters now exhibit clear polynomial behavior. Reflected power remains a scaled version of the detrended SNR, though now in a logarithmic scale, which gives more weight to interference troughs than to peaks. Reflector height and phase shift sensitivities become less distinct, as the heterogeneous weights suppress the signal at high elevation angles. Unfortunately this sensitivity conformation exacerbates the correlation between the two parameters. If one were to design an instrument, uniformly precise SNR measurements (in watts per watt units) would yield more precise reflector height estimates. One region that remains with enhanced discrimination is near the horizon, where reflector height sensitivity vanishes, while the phase-shift sensitivity does not.

Finally, postulating a diagonal covariance matrix is synonymous with assuming uncorrelated measurements. Yet producing a densely populated matrix is memory-demanding to store and time-consuming to process. As a practical alternative, we adopted a sampling spacing that makes it safer to assume independence: we decimated samples regularly spaced in sine of elevation angle (though with eventual gaps). This is preferable because a GPS receiver records measurements regularly spaced in time, which translates into a sparser sampling near the

horizon. Left alone, it would skew the inversion towards the more numerous near-zenith samples, as satellites appear to stall momentarily at culmination point.

The parameters prior covariance matrix $\mathrm{C_x}$ completes the stochastic model specification. It serves to restrict the range of admissible or physically plausible parameter estimates. We have shown above that when the preliminary reflector height is too far off the global minimum, the reflected power bias might inadvertently converge to zero. Fitting a fringe-free SNR trend would lead to a number of undesirable consequences, such as the indeterminacy of reflector height and the singularity of the posterior parameters covariance matrix. We can prevent this numerical adversity by constraining the reflection power to stay within a reasonable deviation of its prior value. This is achieved by setting a smaller value for the reflection power prior standard deviation $\sigma_{B_{\mathrm{dB}}}$ and a large value for all other parameters in the *a priori* variance vector $\Sigma_{\mathrm{x}} = \left[\dots, \sigma^2_{B_{\mathrm{dB}}}, \dots\right]^{\mathrm{T}}$ that forms the diagonal matrix $\mathrm{C_x} = \mathrm{Diag}(\Sigma_{\mathrm{x}})$. Populating off-diagonal elements would be useful were we able to prescribe a certain level of correlation between parameters.

4.5 PERFORMANCE SIMULATIONS

In this section we use simulations to assess the performance of the inversion procedure, in terms of the parameter errors – estimated minus true – and how well they are bounded by the expected parameter uncertainty.

4.5.1 Uncertainty quantification

The (unscaled) parameters posterior covariance matrix,

$$\mathrm{C}_{\hat{\mathrm{x}}} = \left(\mathrm{C}_{\mathrm{x}}^{-1} + \mathrm{J}^{\mathrm{T}}\mathrm{C}_{\mathrm{y}}^{-1}\mathrm{J}\right)^{-1}, \tag{73}$$

is involved in the computation of $\hat{\mathrm{x}}$ itself (eq. (68)) and depends only on the prior specification of the functional and stochastic models, independently from the actual measurements Y themselves. The parameters posterior standard deviation vector $\Sigma_{\hat{\mathrm{x}}} = \sqrt{\mathrm{diag}(\mathrm{C}_{\hat{\mathrm{x}}})}$:

$$\Sigma_{\hat{\mathrm{x}}} = \left[\sigma_{\hat{H}_B}, \sigma_{\hat{\varphi}_B}, \sigma_{\hat{B}_{\mathrm{dB}}^{(0)}}, \sigma_{\hat{B}_{\mathrm{dB}}^{(1)}}, \sigma_{\hat{R}_{\mathrm{dB}}^{(0)}}, \sigma_{\hat{R}_{\mathrm{dB}}^{(1)}}\right]^{\mathrm{T}} \tag{74}$$

is a convenient way to summarize the effect of the large prior matrices $\mathrm{J}, \mathrm{C}_{\mathrm{x}}, \mathrm{C}_{\mathrm{y}}$ (e.g., J has $m \times n$ elements) in terms of just a few scalars.

The uncertainty of parameters depends on the parameters values themselves. For example, in Figure 22 we plot the reflector height uncertainty $\sigma_{\hat{H}_B}$ versus total reflector height $H = H_{\mathrm{A}} - H_{\mathrm{B}}$ and versus reflection power bias B_{dB}. Smaller reflector heights and smaller reflection power produce more uncertain reflector height estimates. This is a direct consequence of the reflector height sensitivity curve becoming confounded with, less distinct from, all other parameters. This is more so as additional parameters are estimated (Figure 23). Furthermore, a diminishing reflection power indicates that interference fringes are vanishing. The reflector height uncertainty also increases with decreasing reflector height because fewer oscillation cycles are included in a track, from horizon to zenith.

The posterior parameters covariance matrix $\hat{\mathrm{C}}_{\hat{\mathrm{x}}} = \hat{\sigma}_0^2 \mathrm{C}_{\hat{\mathrm{x}}}$ is scaled by the variance of unity weight, $\hat{\sigma}_0^2 = SSR/\nu$, where $\nu \approx m - n$ is the statistical degree of freedom, for m elements in the observations vector Y and n elements in the parameters vector X. If functional and stochastic models have been fully specified, $\hat{\sigma}_0^2$ will be close to unity, making this scaling moot. More generally, though, the exact residual variance is not known *a priori* and should be estimated from the sum of squared residuals SSR as above. This scaling contributes to a more realistic parameter uncertainty quantification. It accounts not only for the measurement noise but also for imperfections in the forward model – any mismatch between measurement and simulation (e.g.,

the presence of multiple reflections instead of a single one, as assumed) will lead to an inflated $\hat{\sigma}_0^2$.

4.5.2 Inversion errors

Inversion errors are the discrepancy between retrieved inversion estimates and their true values. Here we focus on the reflector height retrieval. Its inversion error depends primarily on reflector height itself. This is shown in Figure 25. The inversion performance deteriorates with decreasing reflector height, as trend and fringes become confounded. Notice that for a fixed *a priori* height H_A, a smaller total reflector height $H = H_A - H_B$ corresponds to larger unknown height biases H_B, e.g. as when the snow accumulates up to the antenna level (sometimes even beyond it, i.e., snow-buried antenna conditions).

A second factor contributing to reflector height errors is the polynomial degree adopted for the noise power bias K used to detrend observations. Normally we estimate only constant $K_{\mathrm{dB}}^{(0)}$ and linear $K_{\mathrm{dB}}^{(1)}$ terms, which serve to scale the non-linear trend tSNR provided by the forward model. If we could not rely on the aid offered by the forward model and had to detrend observations using solely a higher-order polynomial, this empirical trend would inadvertently take away part of the fringes. Indeed, estimating a quadratic coefficient $K_{\mathrm{dB}}^{(2)}$ causes the reflector height error to soar at small heights. For larger reflector height values detrending is not a challenge.

Inversion error also depends on the reflection power bias. It again exacerbates the chance of confusion between trend and fringes. This is true even assuming noise-less measurements. Random noise further amplifies the reflection power bias effect, as small-amplitude fringes are more easily destroyed; even large-height fringes can be disturbed. We have assumed tracks spanning the full range from horizon to zenith; satellites culminating low in the sky yield more

challenging, shorter tracks. Throughout we have used a fixed 2-m initial reflector height guess (even for smaller reflector heights); a better guess would undoubtedly decrease errors, but it is impractical beyond simulations. We averaged out the dependency on reflection phase-shift. In this case the mean error loses relevance as a near-zero net value might include large positive and negative values; RMS is a more representative statistic.

4.6 DIRECTIONAL DEPENDENCE

4.6.1 Observation importance

One could raise the question of which observations are the most important or contribute the most information. One way of quantifying this is in terms of the impact of each observation on the parameter uncertainty. We formulate this concept starting with the reflector height posterior uncertainty $\hat{\sigma}_{\hat{H}}$ obtained retaining all observations. (The total reflector height uncertainty $\hat{\sigma}_{\hat{H}} = \hat{\sigma}_{\hat{H}_B}$ is simply the reflector height bias uncertainty, as the *a priori* reflector height H_A can be an arbitrary postulated value, thus with $\sigma_{H_A} = 0$.) We withdraw (with replacement) each i-th observation of the total of n observations by removing (one at a time) the corresponding row $\mathrm{J}(i,:)$ from the Jacobian matrix J:

$$\mathrm{J}_i = \begin{bmatrix} \mathrm{J}(1:i-1,:) \\ \mathrm{J}(i+1:n,:) \end{bmatrix} \tag{75}$$

(or rather its modified version $\mathrm{J}' = \mathrm{C}_{\mathrm{y}}^{-1/2}\mathrm{J}$). We utilize the new Jacobian J_i to obtain the respective leave-one-out covariance matrix $\mathrm{C}_{i\hat{\mathrm{x}}} = \left(\mathrm{C}_{\mathrm{x}}^{-1} + \mathrm{J}_i^{\mathrm{T}}\mathrm{J}_i\right)^{-1}$. The observation importance is calculated as the relative increase in the new reflector height uncertainty $\sigma_{i\hat{H}}^2 = \mathrm{C}_{i\hat{\mathrm{x}}}(1,1)$ compared to the original reflector height uncertainty $\sigma_{\hat{H}}^2$:

$$\Delta\sigma_{i\hat{H}}^2 = \left(\sigma_{i\hat{H}}^2 - \sigma_{\hat{H}}^2\right)/\sigma_{i\hat{H}}^2 \tag{76}$$

which is guaranteed to be non-negative (assuming uncorrelated and equal-variance observations). This implies that the parameter uncertainty cannot be decreased with fewer observations; conversely, including more observations cannot increase the parameter uncertainty. (This statement is applicable in the average sense, not necessarily to individual realizations, as the scaling $\hat{\sigma}_0^2$ is neglected.) Yet the presence or absence of observations at certain locations is inconsequential as far as the posterior parameters uncertainty is concerned, while others are crucial.

If we were estimating reflector height only, the observation importance would be proportional to the respective observation sensitivity, squared. (Below we consider multiple simultaneously estimated parameters.) This relationship can be demonstrated as follows. The inner product $\mathrm{J}^{\mathrm{T}}\mathrm{J} = N$ for a single-column Jacobian J is a sum $N = \sum_{k=1}^{n} N(k)$ of n squared sensitivities $N(k) = J^2(k)$. Assuming no prior covariance, its posterior counterpart is simply $\mathrm{C} = N^{-1}$, which contains just the variance, $\mathrm{C} = \sigma^2$. The leave-one-out case is $\sigma_i^2 = \mathrm{C}_i = N_i^{-1}$, where $N_i = N - N(i)$ is a sum $N_i = \sum_{k=1}^{i-1} N(k) + \sum_{k=i+1}^{n} N(k)$ that leaves the i-th term out. The observation importance can then be expressed as:

$$\Delta\sigma_i^2 = (\sigma_i^2 - \sigma^2)/\sigma_i^2 = (N_i^{-1} - N^{-1})N_i = N(i)N^{-1} = J^2(i)\sigma^2, \tag{77}$$

which is $\Delta\sigma_i^2 \propto J^2(i)$ since $\sigma^2 =$ const., as claimed. The resulting observation importance can be envisioned as an unsigned version of the observation sensitivity with respect to reflector height show in Figure 21 (top-most blue trace).

Observation importance and sensitivity exhibit peaks and troughs related to the conditions at hand, i.e., parameter values $\mathrm{X} = [H_B, \varphi_B, \ldots]^{\mathrm{T}}$ in $\sigma_{\hat{H}}^2\big|_{\mathrm{X}}$. A simplification is found in the mean importance:

$$\Delta\bar{\sigma}^2_{i\hat{H}} \approx 4^{-1}\sum_{k=1}^{4}\Delta\sigma^2_{\hat{H}}\big|_{\varphi_B(k)} \tag{78}$$

where the averaging takes place over a regularly spaced phase-shift domain $\varphi_B = [0, 90°, 180°, 270°]^{\mathrm{T}}$ (or more than four values, if necessary). Figure 24 shows mean importance as a smooth thick gray line; the contributing phase-dependent individual importances are shown as thin gray lines in the background. The former is essentially the envelope of the latter ones.

In practice, reflector height is not the only parameter being estimated. In this case, observation importance retains its definition, yet it exhibits a more complex behavior. Its interpretation is complicated by the fact that the information contributed by each observation is not used solely to determine reflector height, rather is dispersed among all parameters. In other words, the sensitivities with respect to each parameter become intertwined. In Figure 24 a thick light-blue line is used to represent the observation importance in determining reflector height when a phase shift is also estimated, $\Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=2}$ (two simultaneous parameters). Notice the curve has shifted to the right compared to when estimating a single parameter, $\Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=1}$; observations at higher elevation angles become more important. Also the near-horizon region gains in importance; this is a manifestation of the enhanced discrimination, between reflector height and phase-shift sensitivities, as discussed for Figure 20. When the number of simultaneous parameters is raised up to the level used in practice, $\Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=6}$, the observation importance becomes more detailed (light-red thick curve), although the overall shape follows the previous one (light-blue thick curve), $\Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=2}$.

Results are dependent not only on the number of parameters but also on their values. In the discussion above we adopted a 1-m reflector height. In the same Figure 24, thin darker lines

represent results lowering reflector height to $1/3$ m. Observation importance remains largely unaffected if were we estimating height only (thin black line), i.e., $\Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=1}_{H=1/3\mathrm{m}} \approx \Delta\bar{\sigma}^2_{\hat{H}}\big|^{p=1}_{H=1\mathrm{m}}$. Yet when actually estimating multiple parameters (thin dark-red and dark-blue lines), the curves become even broader, less peaked. Observations at the highest elevation angles become important for a successful separation of SNR trend and fringes for small heights, $\Delta\bar{\sigma}^2_{\hat{H}}(e = 90°)\big|^{p=6}_{H=1/3\mathrm{m}} \gg \Delta\bar{\sigma}^2_{\hat{H}}(e = 90°)\big|^{p=1}_{H=1\mathrm{m}}$.

4.6.2 Parameter reach

A related question is at what elevation angles the parameter estimates are most relevant or best determined. Here we focus on the phase function parameters instead of reflection or noise power parameters.

We can utilize the estimated reflector height $\widehat{H}_B$ and phase-shift $\hat{\varphi}_B$ to evaluate the full phase bias function $\hat{\phi}_B = \hat{\varphi}_B + 4\pi\lambda^{-1}\widehat{H}_B \sin e$ over varying elevation angles e. Similarly, we can extract the corresponding 2-by-2 portion of the parameters posterior covariance matrix $\hat{\mathrm{C}}_{\hat{\mathbf{x}}}$, containing the uncertainty for reflector height $\hat{\sigma}_{\hat{H}_B}$ and for phase-shift $\hat{\sigma}^2_{\hat{\varphi}_B}$, as well as their correlation ρ,

$$\begin{bmatrix} \hat{\sigma}^2_{\hat{H}_B} & \rho\,\hat{\sigma}_{\hat{H}_B}\hat{\sigma}_{\hat{\varphi}_B} \\ \rho\,\hat{\sigma}_{\hat{H}_B}\hat{\sigma}_{\hat{\varphi}_B} & \hat{\sigma}^2_{\hat{\varphi}_B} \end{bmatrix}, \tag{79}$$

which is then propagated to obtain the evaluated phase function uncertainty $\hat{\sigma}_{\hat{\phi}_B}$, see Figure 26.

The uncertainty $\hat{\sigma}_{\hat{\phi}_B}$ attains a clear minimum versus elevation angle. The least-uncertainty elevation angle $e_p = \operatorname{argmin}\hat{\sigma}_{\hat{\phi}_B}$ can be found through $\partial\hat{\sigma}_{\hat{\phi}_B}/\partial e_p = 0$, whose solution reads:

$$4\pi\lambda^{-1}\sin e_p = -\frac{\hat{\sigma}_{\hat{H}_B} + \hat{\sigma}_{\hat{\phi}_B}}{2\rho\hat{\sigma}_{\hat{H}_B}}. \tag{80}$$

We will call e_p peak elevation angle as it pinpoints the observation direction where reflector height and phase-shift are best determined. The azimuth and epoch coinciding with the peak elevation angle act as track tags, later used for clustering similar tracks and analyzing their time series of retrievals.

Finally, if we normalize phase uncertainty by its value at the peak elevation angle, $\hat{\sigma}_{\hat{\phi}_B}(e)/\hat{\sigma}_{\hat{\phi}_B}(e_p)$, then plot such relative weights versus the radial distance to the center of the first Fresnel zone at each elevation angle, we obtain Figure 27. It can be interpreted as the reflection footprint.

4.7 CONCLUSIONS

We have formulated a forward/inverse approach for the estimation of snow depth from GPS SNR observations. After briefly reviewing the multipath forward model, we expanded it to include changing environmental conditions such as snow depth. Then we described the inverse model in terms of its functional and stochastic components. Throughout, we used simulations to illustrate various aspects, such as trend and fringes; sensitivity of observations to parameter changes; the parameter space in which the objective function is embedded (including the indeterminacy of reflector heights under small reflection amplitude conditions); and the expected reflector height inversion error (i.e., retrieved minus true), and how well it was bounded by the model uncertainty. In part II, we apply the model to measurements collected at three different locations over a multi-year period.

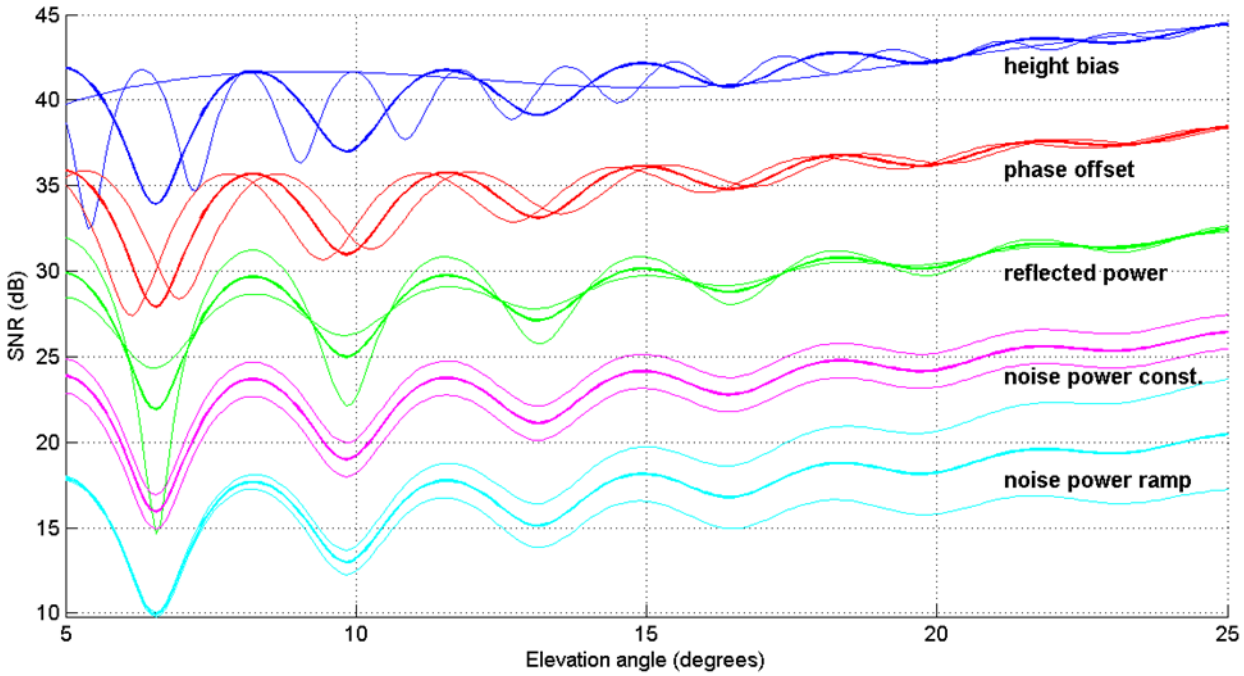

Figure 16: Effect of each parameter on SNR observations; curves are displaced vertically for clarity.

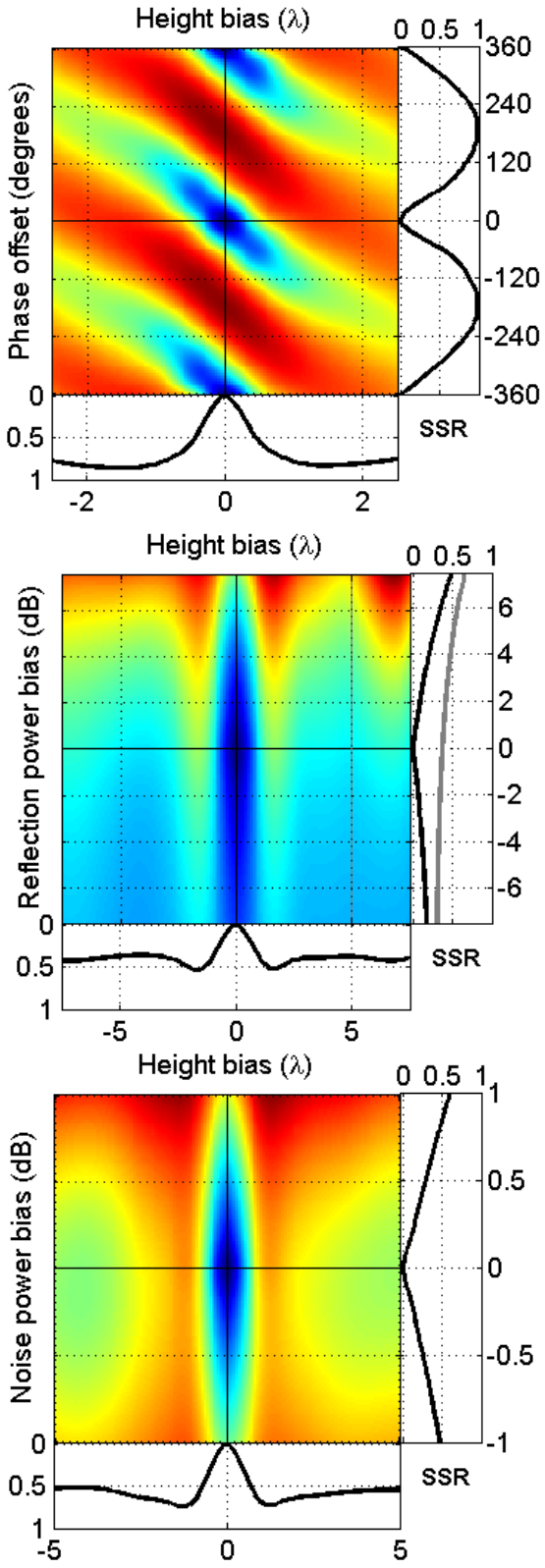

Figure 17: Minimization objective function (sum of squared residuals, SSR) over the parameter space over which the inversion takes place; values are normalized to their peak value at each panel.

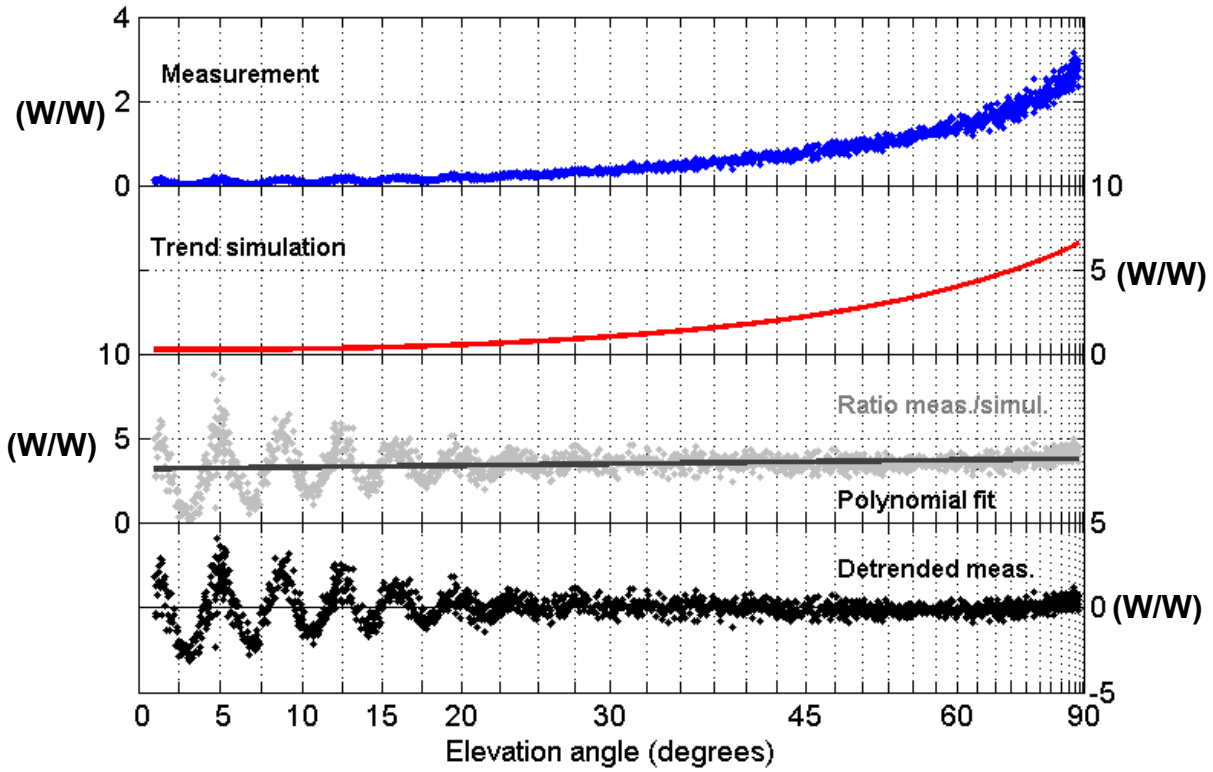

Figure 18: Noise power bias bootstrapping involved in the preliminary detrending; each panel has been scaled separately for clarity (original units, W/W, are normalized by an arbitrary constant for clarity).

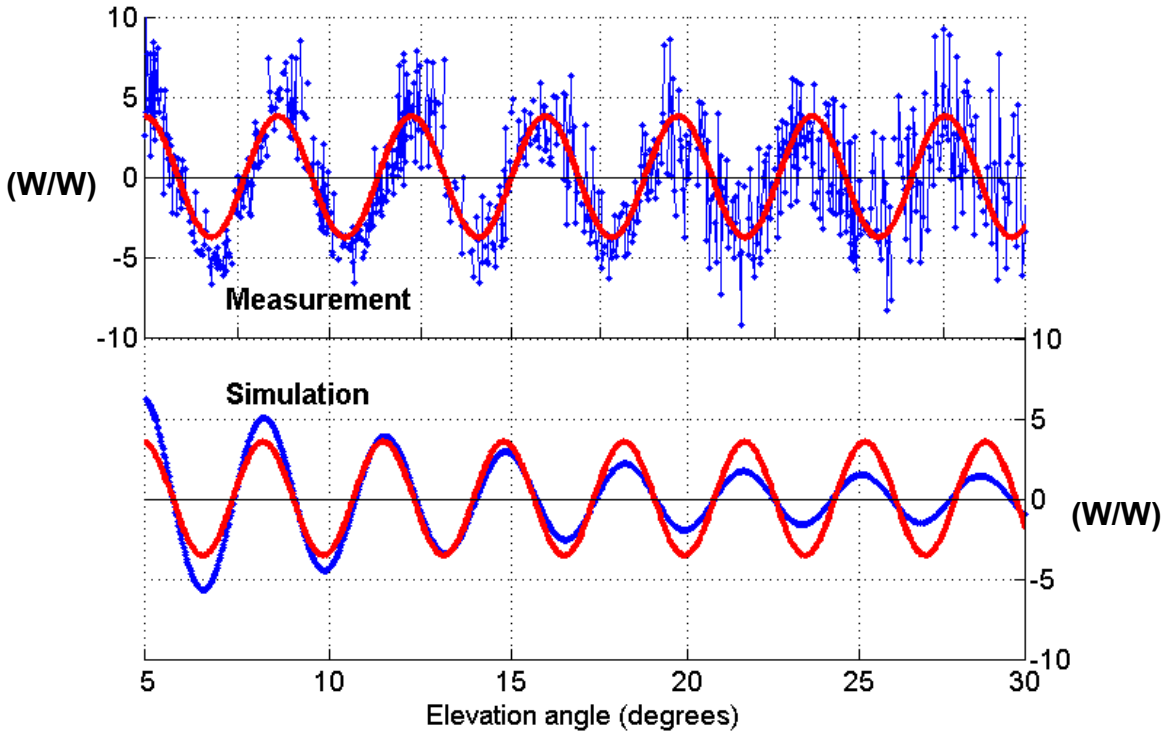

Figure 19: Reflection bias bootstrapping involved in the preliminary fringe fitting; each panel has been scaled separately for clarity (original units, W/W, are normalized by an arbitrary constant for clarity).

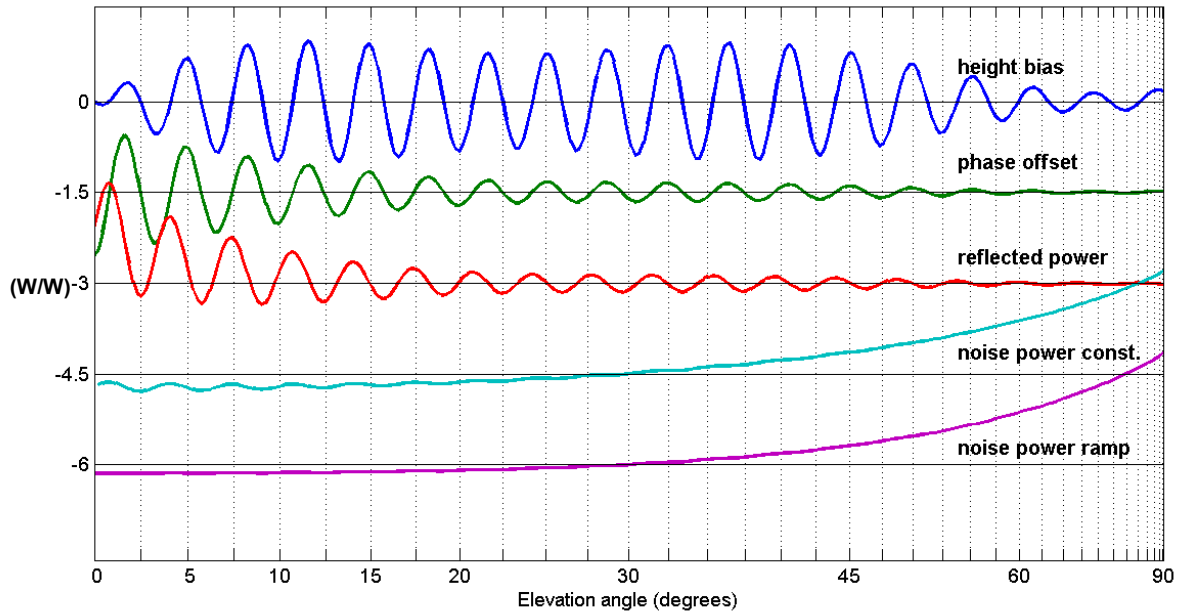

Figure 20: Sensitivity of observations (in original units of W/W) to changes in each bias; curves are scaled and displaced vertically for clarity.

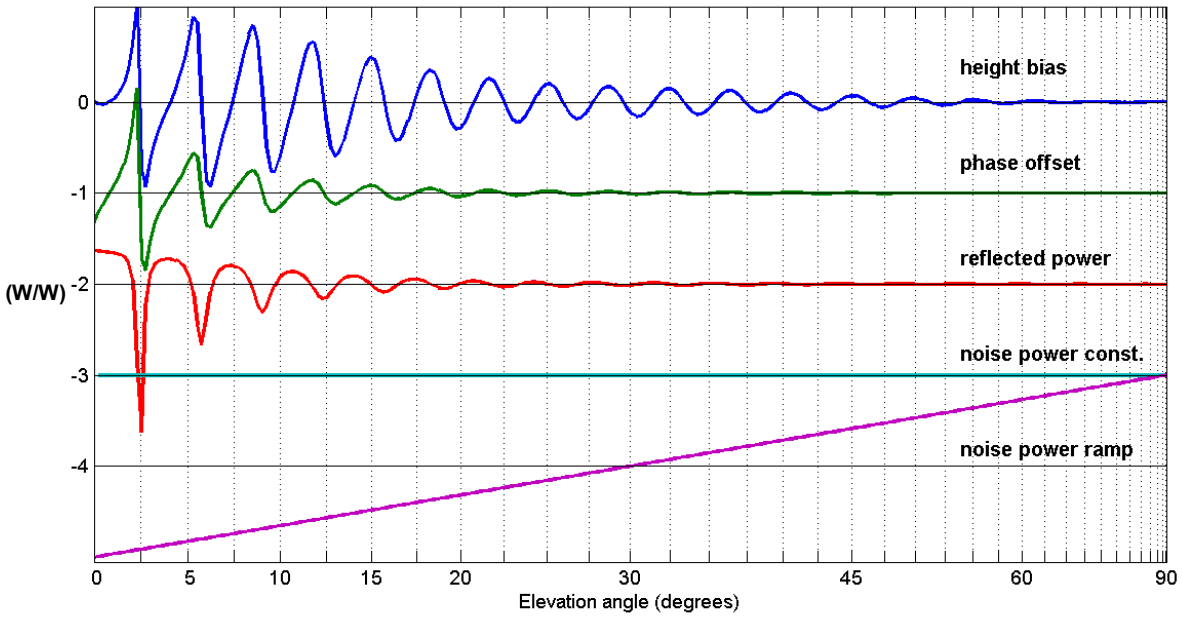

Figure 21: Sensitivity of observations (in original units of decibels) to changes in each bias, after accounting for the varying observation uncertainty; curves are scaled and displaced vertically for clarity.

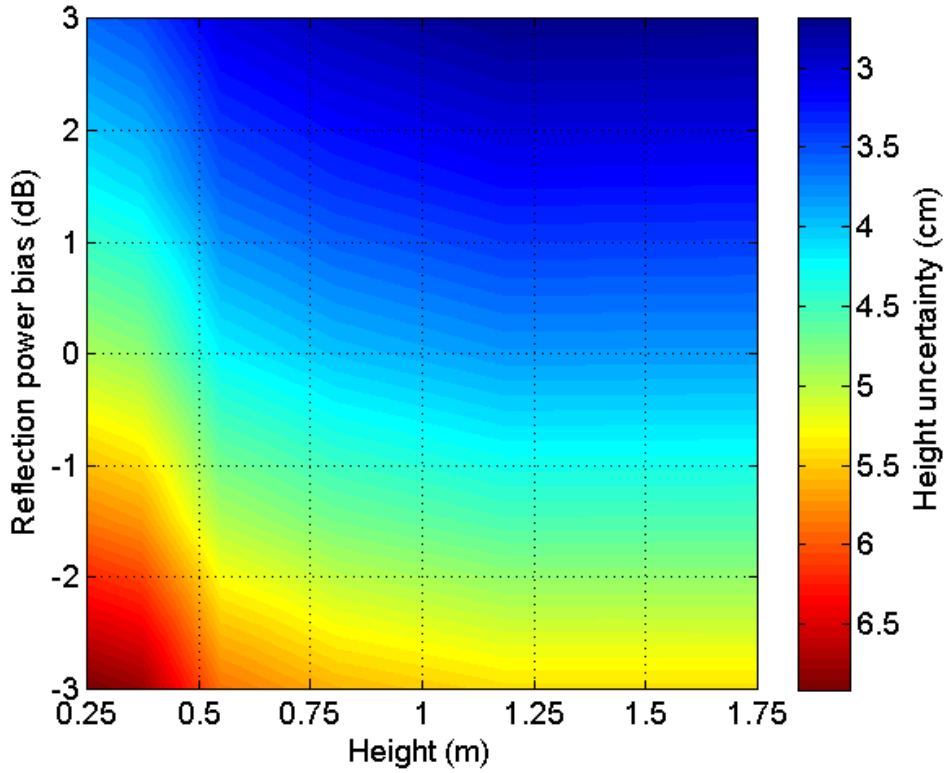

Figure 22: Reflector height uncertainty versus total height and reflection power bias.

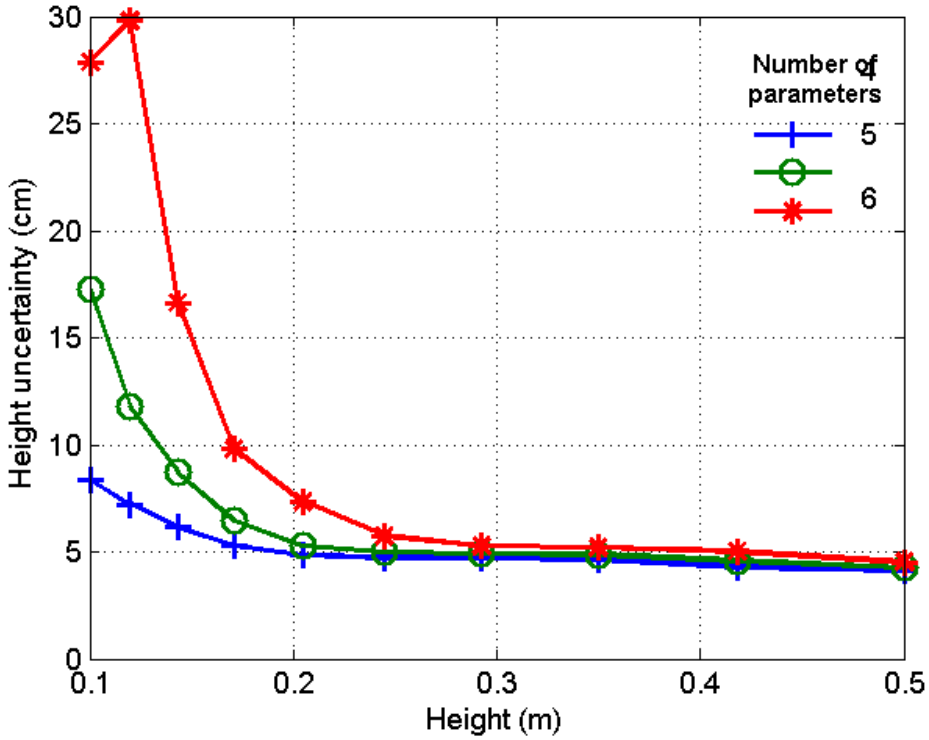

Figure 23: Reflector height uncertainty versus total reflector height, over increasing number of parameters.

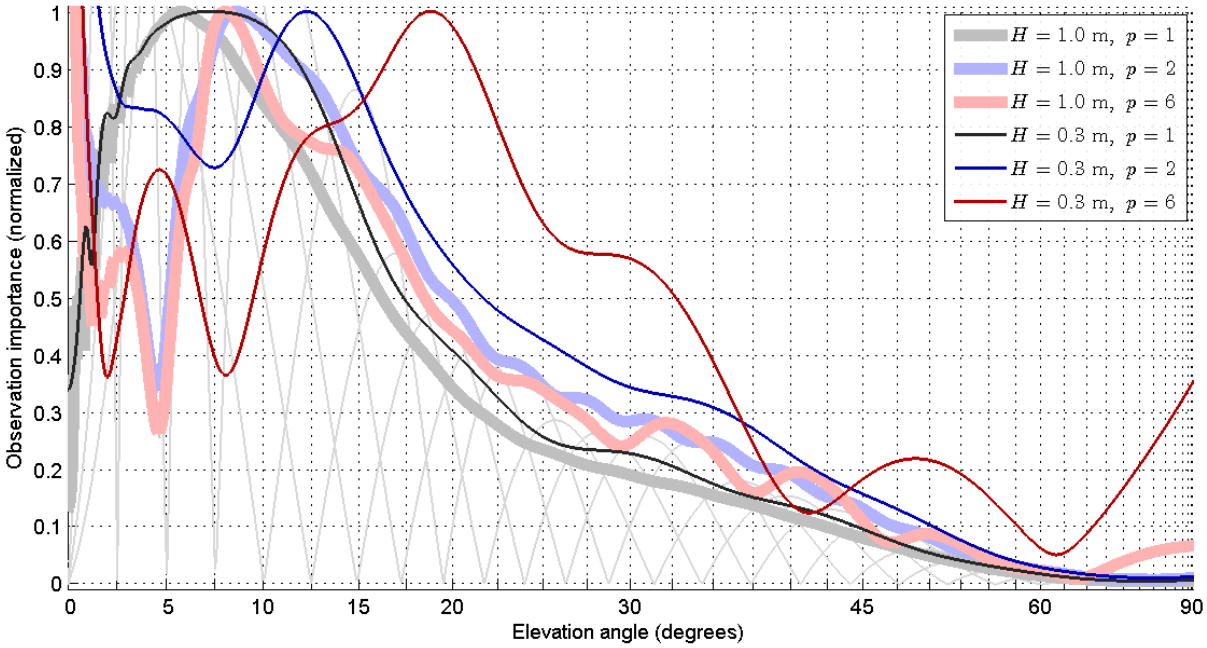

Figure 24: Observation importance in determining reflector height under different conditions: total reflector height value and number of simultaneous parameters.

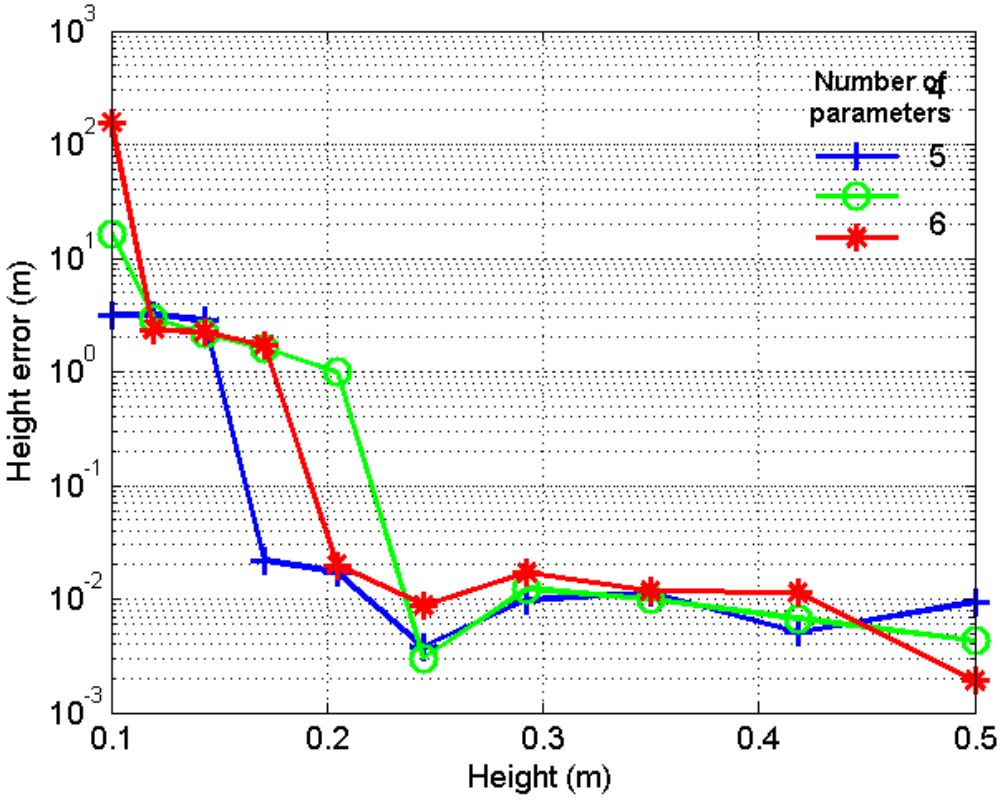

Figure 25: Reflector height error (estimated minus true) versus true reflector height.

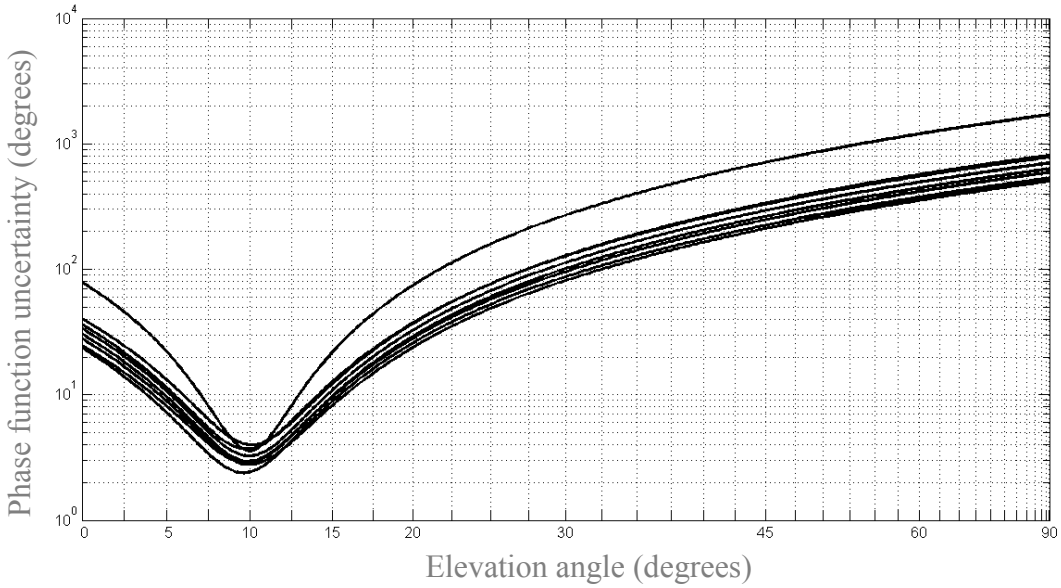

Figure 26: Formal uncertainty of phase function.

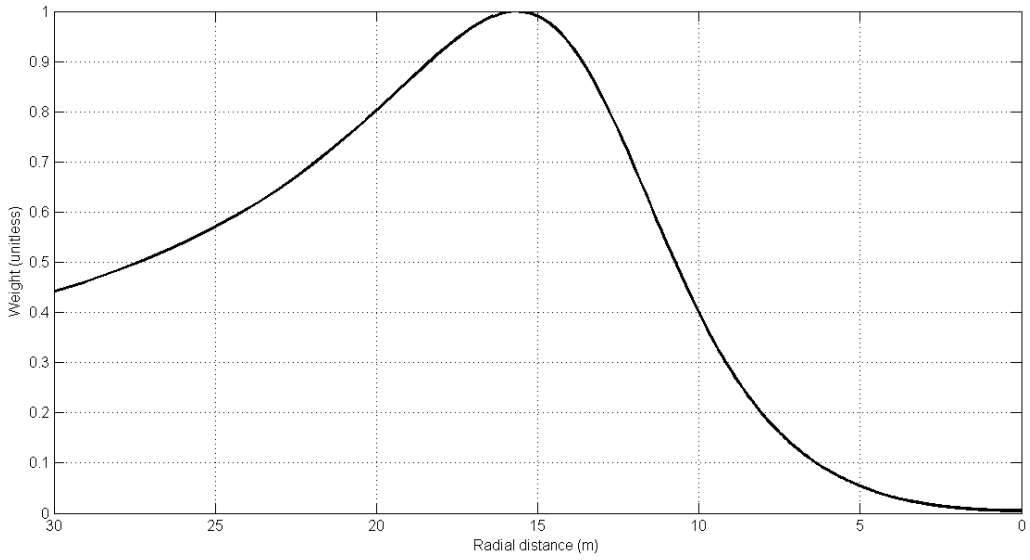

Figure 27: Sensing footprint.

Chapter 5: Inverse modeling of GPS multipath for snow depth estimation – Part II: Demonstration and assessment

5.1 INTRODUCTION

This is part II of a contribution that uses a statistical inverse model for GPS multipath to estimate snow depth. In part I we presented the model formulation and illustrated its main features using simulations. In part II we apply the model to GPS observations collected at three sites over a multi-year period for retrieving snow depth. These sites were chosen because they represent a variety of conditions and levels of difficulty, including a flat grassland site (simple but with limited *in situ* data); an alpine site (with significant terrain, with multiple years of *in situ* data); and a forested site (most difficult, with some *in situ* data). We start with a general development applicable to all sites, in which intermediary results are explored in more detail. Then we proceed to show the final results – namely, snow depth time series – at each of the three sites, validating them against independent *in situ* measurements.

The fit of SNR observations described in part I provides parameter estimates and their covariance matrix, as well as observation residuals, for each satellite track. Here we analyze the resulting parameters and residuals. We start examining a few representative fits, illustrating and discussing the origin of a variety of good and bad conditions, such as measurement noise, well- vs. poorly-determined reflector heights, instrument-related issues, etc. Then we discuss a methodology to quality control these estimates based on track clusters; the thousands of tracks retrieved in a year can thus be analyzed in terms of only 10 to 20 units. We introduce a specially designed diagram as a convenient summary of the track clusters available in a site. Such a repeatable sensing configuration allows us to compare tracks belonging to the same cluster with the purpose of detecting and rejecting anomalous conditions, as physically plausible

environmental changes do not happen in isolation over time and space. This principle leads to a number of strategies for quality control of results, which is demanded for operational use of GPS snow sensing. Track clusters, as statistically homogeneous units, play a critical role in in quality control as outliers are only defined in comparison to the tendency and dispersion of results. Daily site averages are then compiled from the results. As a pre-requisite, the different clusters must be harmonized, to ensure that noise is filtered out and the signal of interest is enhanced rather than suppressed. This means accounting for genuine azimuthal asymmetry in the distribution of snow around the antenna, and also dealing with issues such as assigning statistical weights to varying quality track clusters (e.g., subject or not to partial obstructions in sky visibility).

5.2 General development

In this section we examine the matching of model and measurements, in terms of observation residuals; the satellite coverage, over time and space; the quality control procedure applied to mitigate anomalous results on a cluster-by-cluster basis; and combinations of estimates across different track clusters.

5.2.1 Observations

Here we compare and contrast measured and modeled GPS SNR observations through varying conditions. A typical good fit is shown in Figure 28, corresponding to the beginning of the snow season, at which snow height is increasing. The fit of the model developed in part I to SNR measurements is affected by a number of factors. Here we discuss aspects that degrade the fit but do not necessarily cause a bias in reflector height estimates.

5.2.1.1 Secondary reflections

Throughout we have assumed the existence of a single specular reflection, which matches large and planar surfaces. Truncated and/or non-planar surfaces represent a departure from this

assumption; the first case transforms a simpler reflection into a more complicated diffraction phenomenon, while the second case might as well introduce secondary reflections, originating from disjoint surface regions. Interference fringes become convoluted with multiple superimposed beats, as instead of a single interferometric phasor, there is an additional one (between direct and secondary reflection) and yet another one (between the two reflections). As long as there is there is a unique dominating reflection, the inversion will have no difficulty fitting it, as the extra reflections, however structured, will remain zero-mean (see Figure 29). Contrarily, two reflections having comparable strength will be poorly fit by the present model.

5.2.1.2 Interferometric power effects

Random deviations of the actual surface with respect to its undulated approximation – called roughness or residual surface height – will affect the interferometric power, P_{i}. SNR measurements will exhibit a diminishing number of significant interference fringes, compared to the measurement noise level. Although having fewer fringes (Figure 30) facilitates the model fit, the reflector height parameter may become ill-determined – its estimates will be more uncertain.

Surface roughness varies from larger to smaller values for new pristine snowflakes, older settled snow, and finally melting conditions. These changes in the snowpack surface are accompanied by corresponding changes in the snowpack volume. Very similar to the roughness effect, changes in density affect fringe amplitudes. So care needs to be exercised in any future attempt at assigning a physical interpretation to reflection power bias retrievals B_{dB}. Finally, as the scattering medium transitions at the beginning and end of the season, between more homogeneous snow and soil, the model/measurement fit generally deteriorates, particularly when the mixture involves partially covered vegetation. In fact, even *in situ* manual measurements of

snow depth are problematic in those situations – e.g., 5-cm deep snow patches littering 50% of the ground area must not be mistaken for a uniform 2.5-cm thick layer of snow.

5.2.1.3 Direct power effects

Snow precipitation can affect SNR measurements through the direct power P_{d}. Precipitation of various form are known to attenuate satellite-to-ground radio links. First, this shifts the SNR measurements up or down (in decibels). Second, it tilts the SNR trend, $\mathrm{tSNR} = (P_{\mathrm{d}} + P_{\mathrm{r}})P_{\mathrm{n}}^{-1}$, as atmospheric attenuation is least for a satellite at zenith and increases monotonically with decreasing elevation angle. Third, SNR fringes $\mathrm{dSNR} = 2\sqrt{P_{\mathrm{d}}^{\mathrm{C}}}\sqrt{P_{\mathrm{r}}^{\mathrm{C}}}P_{\mathrm{n}}^{-1}\cos\phi_{\mathrm{i}}$ will change in amplitude because of the decrease in the coherent component of the direct power, $P_{\mathrm{d}}^{\mathrm{C}}$ (this is in contrast to the trend tSNR, which involves both components of the direct power, $P_{\mathrm{d}} = P_{\mathrm{d}}^{\mathrm{C}} + P_{\mathrm{d}}^{\mathrm{I}}$, coherent and incoherent). Ideally the reflection power bias B_{dB} would isolate exclusively the interferometric power effects, but it may end up contaminated by unaccounted for direct power variations – which is yet another reason for refraining from ascribing a single physical origin to B_{dB}.

Partial obstructions can affect either or both direct and interferometric powers. SNR measurements, albeit corrupted, are still recorded. This situation is unlike complete blockages, as caused by, e.g., topography. The deposition of snow and the formation of a winter rime on the antenna are a particularly insidious type obstruction, as their presence in the near-field of the antenna element can easily distort the gain pattern in a significant manner. Because snow accumulates more easily on top of the antenna, the impact is more common on the satellite direction rather than on the reflection, which has incidence at negative elevation angles. In the far-field, trees are another important nuisance, so much so that their absence is held as a strong requirement for the proper functioning of multipath-based reflectometry (Larson and Nievinski

2013). Trees can constitute an obstruction or a blockage, depending on the distance from the antenna and also on the time of the year in the case of seasonal plants, as they put forth foliage then shed their leaves.

5.2.1.4 Instrument-related effects

Measurements are a manifestation not only of genuine physical effect as above, but also more mundane spurious instrumental effects. It is importance to disentangle the two, as the former can be of interest while the latter ones are undesirable. For example, one should expect satellite-specific direct power offsets and, in the long-term, also power drifts, as spacecrafts age and modernized designs are launched. Less subtle incidents are the early termination in the recording of a rising or setting satellite, a consequence of maxing out the finite number of tracking channels available in the receiver, which are allocated giving greater priority to higher-elevation satellites. As another example, noise power depends on the state of conservation of receiver cables, and also on their physical temperature. Finally, it is important to recall that the L2C code is actually two, a data-modulated medium-length sub-code (CM) and a data-less longer one (CL); CL has a 3-dB effective power advantage over CM (Fontana et al. 2001). SNR occasionally exhibits steps (Figure 31), ~ 3-dB high and approximately regular width. Such a sudden and sharp step up and down is physically implausible; rather, the receiver appears to be switching between tracking the two sub-codes, presumably making the decision based on a threshold exceeded.

As we have seen in this section, anomalous conditions may result in measurement spikes, jumps, and short-lived rapidly-varying fluctuations. The ones with a physical origin might be of interest in themselves, whereas those related to faulty instrumentation are obviously not. In any case, for snow depth sensing purposes, it is necessary and sufficient to be able to either neutralize

such measurement outliers through a statistically robust fit or detect suspicious fits and discard the problematic ones that could not otherwise be salvaged. This issue is addressed in the next second section below. First we need to consider the satellite coverage.

5.2.2 Satellite coverage and track clustering

As stated in part I, we have used the newer L2C signal. Of the thirty one GPS satellites, ten currently transmit L2C. Satellite observations are partitioned into rising and setting portions (as discussed in part I), yielding approximately 20 unique track clusters at station P360. Figure 33 shows track azimuths over a year at P360. It demonstrates that all quadrants are adequately sampled at this particular sit, which is to say it has good visibility conditions. It also highlights the repeatability of GPS orbits, with deviations at the few-degree range over that period; this translates into ~ 50-100 cm azimuthal displacement in the first Fresnel zone at 10-15° elevation angle, assuming a 2-m high antenna. This repeatability facilitates the separation of retrieved reflector heights into their ground- and snow-related components, as it allows assuming the former to be constant. Galileo, the European equivalent of GPS, will not have sidereal repeatability; neither does Glonass. Tracks can thus be clustered by azimuth.

For a given track cluster, its revisit time is also repeatable, amounting to practically one sidereal day. The deficit in time relative to a calendar day results in the track time of the day receding ~ 4 min. every day. This slow but steady accumulation eventually makes the time of day to return to its starting value after approximately one year (Figure 34). As all GPS satellites drift similarly (though not exactly at the same rate), the time between successive tracks remains nearly repeatable. Its reciprocal, the sampling rate, has median equal to approximately one track per hour, with a low value of one track within two hours and a high of one track within 15 min; both extremes occur every day, with low-rate idle periods in interstices with high-rate bursts.

Anomalous rates are caused by receiver failure. The time of the day reduced to a fixed day could similarly be used to cluster tracks. Neighboring track clusters, that are close in azimuth and/or in reduced time of the day are expected to be more comparable, as they sample more similar conditions and are subject to similar errors.

5.2.3 Quality control (intra-clusters)

The key to quality control (QC) is in grouping results into statistically homogeneous units, i.e., having measurements collected under comparable conditions. In our case, azimuth-clustered tracks are the natural starting unit. Secondarily, we must employ a running average to account for varying quality of results, from beginning to peak to the end of the snow season. The detection of anomalous results further requires an estimate of the statistical dispersion to be expected. Considering that the sample is contaminated with outliers, robust estimators – median instead of the mean, and median absolute deviation over the standard deviation – are called for, if the first- and second-order statistical moments are to be representative. Given these two estimates, a tolerance interval can then be constructed such that it bounds, say, a 99% proportion of the valid results with 95% confidence level. Unfortunately sometimes we will end up rejecting fits which turn out to be false positives; a certain number of such incidents are bound to happen at a rate equal to the significance level corresponding to the complement of the confidence level above, i.e., 5%. Yet we wish quality control to be judicious, or else too many valid estimates will be lost.

One aspect that requires further care is with regard to statistics that involve a sum of squared values, as they are expected to follow a chi-squared distribution. To accommodate this requirement, we handle their values in logarithm form, which guarantees that tolerances remain non-negative while also transforming the probability distribution to restore normality. After

applying that transformation, we can calculate the moving average and dispersion as usual and then form the desired tolerance bound, which is converted back into original units via the exponential. Contrary to the two-tailed tolerance intervals that are adequate for first-order statistics, a single-tailed tolerance bound was found more appropriate for second-order statistics, as a small dispersion is not to be avoided.

No single statistic detects all the outliers. Ideally we would avoid directly handling reflector heights themselves in doing this QC, as it would risk imposing our expectation on the results – instead of letting the data speak for themselves. In the sub-sections that follow we demonstrate four particular indirect statistics found to be useful; their application is illustrated using a recurrent track cluster. The corresponding time series of estimated reflector heights is shown in Figure 35 (top panel); these raw reflector heights will be transformed into snow depth later. Even in raw form, the snow season can be clearly distinguished by the dramatic increase in reflector height. Several snow precipitation events are discernible as sudden rises in reflector height, followed by a slower decay indicating snow settling. At the end of the season, we find unabated melting, until bare ground eventually becomes exposed. Over the remainder of the year, reflector heights varies just a few centimeters at this site, which indicates that our *a priori* value for the height of the antenna above the ground was accurate and also that the ground is nearly horizontal and mostly devoid of tall/dense vegetation.

5.2.3.1 Degree of freedom

The simplest statistic is the degree of freedom, essentially the number of observations per track (modulo a constant number of parameters). The time series shown in Figure 35 (bottom panel) is very stable, so a global low-order polynomial suffices for this statistic. It does a good job detecting problems related to data outages.

5.2.3.2 Goodness of fit

We use the regression root-mean-square of residuals, $\hat{\sigma}_0$ (scaled to account for the multiple parameters) to test the goodness-of-fit. Figure 35 (second panel from top) shows a time series of $\hat{\sigma}_0$ and its tolerance interval. We succeeded in neutralizing the risk posed by outliers in distorting the tolerance intervals. Tracks near days -25 reject the null hypothesis of statistically equivalence against the long-term average. Notice this statistic is non-stationary, with best fits being more common during the period when snow depth is relatively stable, between days 100 and 125.

5.2.3.3 Reflector height uncertainty

Sometimes the fit is good but the reflector height uncertainty is bad as the signal is evidently gone – there are not many fringes above the measurement noise level (Figure 30). Such cases go undetected by the previous statistic, but fortunately many are picked up by the present one. Low-uncertainty fits are common during heavy snowfalls.

5.2.3.4 Peak elevation angle

The peak elevation angle (defined in the part-I paper) behaves very much like a random variable, as it is determined by a multitude of non-dominating factors. It is straightforward then to form tolerance intervals and use these to detect outliers. This statistic was found to perform especially well in cases that were particularly challenging for the previous statistics. For example, some fits yielded small residuals and produced unsuspecting reflector height uncertainties, but the result corresponded to a significantly different peak elevation angle. This was the case not just because the reflection signal that is typically found in the measurements turned out to be missing on that day. The aggravating factor is the presence of interference fringes at non-typical elevation angles (Figure 32). Despite being well fit and well determined, they are generated by different reflecting conditions, compared to those found on most other days.

5.2.4 Combinations (inter-clusters)

In the previous section we dealt with results on a cluster-by-cluster basis. Now we are interested in combining multiple clusters. The main purpose is to average-out random noise. Mitigating noise aims at not only coping with measurement errors but also compensating for model deficiencies, to the extent that they are not in common across different clusters.

5.2.4.1 Vertical datum

Before we combine different clusters, we have to address their long-term differences. More specifically, we discuss how best to convert retrieved reflector heights into physically-meaningful snow thickness and ultimately snow depth, the latter being more comparable across different track clusters (Figure 37). We wish to remove the temporally permanent spatially non-uniform component of snow sensed at different track clusters.

Given cluster-wise time series of estimated reflector heights biases H_{ij} – for simplicity we drop the original subscript B in favor of i and j for each day and cluster, respectively –, their weighted median over the snow-free period is taken as the ground height, constant on a cluster-by-cluster basis:

$$\bar{H}_j = \underset{i \notin \mathcal{W}}{\text{med}}\, H_{ij} \tag{81}$$

Weights follow from the track inversion uncertainties. The ground height uncertainty is a consequence of these uncertainties as well as the day-to-day reflector height dispersion. As the snow-free period typically spans several months, the median reflector height can be determined down to the millimeter level.

In GPS-MR, reflector heights are reckoned from the horizontal plane passing through the antenna phase center (Figure 36); upon the adoption of a non-zero *a priori* reflector height, this plane is effectively displaced vertically, lowered down to the intersection of the antenna mount

pole with the ground. This type of vertical coordinate is unsuitable for snow sensing as it is affected by large-scale ground topography – snow surface heights will be smaller downhill and greater uphill. Subtracting ground heights from their respective snow surface heights results in snow thickness values,

$$T_{ij} = H_{ij} - \bar{H}_j \tag{82}$$

which is a completely physically unambiguous quantity. Snow thickness is more comparable than snow heights across varying-azimuth track clusters.

Although the largest-scale ground topography has been successfully isolated, further uniformity can be achieved because an avoidable smaller-scale component persists. Notice that snow thickness variability is impacted the underlying ground surface, even when the overlying snow surface is relatively uniform. This spatial non-uniformity can be easily estimated postulating it to be temporally permanent. Snow depth is then defined here as

$$D_{ij} = T_{ij} - \delta\bar{T}_j \tag{83}$$

by further isolating the cluster-wise deviation in median thickness,

$$\delta\bar{T}_j = \bar{T}_j - \bar{\bar{T}} \tag{84}$$

i.e., removing the cluster-wise median thickness

$$\bar{T}_j = \underset{i \in \mathcal{W}}{\text{med}}\, T_{ij} \tag{85}$$

and restoring the site median thickness

$$\bar{\bar{T}} = \underset{j}{\text{med}}\, \bar{T}_j \tag{86}$$

The interpretation of a topographical origin for the cluster-wise deviation in median snow thickness $\delta\bar{T}_j$ is corroborated by its good correlation with the cluster-wise deviation in median ground height, $\delta\bar{H}_j = \bar{H}_j - \bar{\bar{H}}$, where the $\bar{\bar{H}} = \text{med}_j\, \bar{H}_j$ (see Figure 38) – in spite of the two being

calculated over disjoint sets (snow-covered period $i \in \mathcal{W}$ in the case of $\bar{T}_j$ and snow-free period $i \notin \mathcal{W}$ for $\bar{H}_j$). As these deviations $\delta\bar{T}_j$ sum up to zero, average depth can continue to be interpreted similarly to average thickness.

5.2.4.2 Averaging

The averaging of snow depths collected at different track clusters must take into consideration the varying uncertainty of each contributing track. Although the least squares procedure does provide uncertainties, they need to be scaled – this is an instance of the so-called variance factor problem.

A solution is achieved in three steps. It starts with a preliminary weighted running median D_k, calculated at a given spacing $k = 1,2, \ldots$ (say, daily postings), with overlapping or not. The preliminary post-fit residuals $\delta D_{ij} = D_{ij} - D_k$ then go through their own averaging – necessarily employing a wider averaging window (say, monthly) –, which produces scaling factors for the original least squares uncertainties. The running weighted median is then repeated, producing final averages. These three steps can optionally be repeated until convergence; after the first iteration they are already closer to unity (Figure 39). Inspecting the scaling factor, it is readily apparent that some clusters are better than others, while other clusters are consistently worse than most; in fact, some clusters may very well not be worth processing at all, but that is not something generally known *a priori*.

Averaging is also an additional opportunity for quality control – a tolerance interval can be built, as before. Such an inter-cluster QC picks up outliers that went undetected through the previous intra-cluster indirect QC. The cleaned sample can then be employed to redo the averaging, via a weighted mean instead of the weighted median. Although the mean is disastrous under the possibility of outliers, when the sample is guaranteed to be approximately normally

distributed, it has the greatest statistical efficiency – its variance is ~ 50% smaller than that of the median.

5.3 SITE-SPECIFIC RESULTS

This methodology is now further explored at three stations (Table 1) and over a longer period (up to 3 years). Throughout, we assess the performance of the GPS against independent co-located *in situ* measurements (Table 2). We also compare the GPS estimates to the nearest SNOTEL station. Although not co-located with GPS, SNOTEL data – widely used for operational snow monitoring in the U.S – are important because they provide accurate information on snowfall events.

Table 1: Site coordinates.

Environment	Code	Latitude (degrees)	Longitude (degrees)	Altitude (m)
Grassland	P360	44.317852	-111.450677	1857.9
Forested	RN86	41.864856	-111.502514	2590.7
Alpine	NWOT	40.055387	-105.590527	3522.5

Table 2: *In situ* data quantity.

Environment	Duration	Interval	# Epochs	Type	Replication
Grasslands	6 mon.	~ 6 h	500	webcam	1 value/epoch
Forested	7 mon.	~ 2-3 week	9	manual	20-150 values/epoch
Alpine	3 yr.	~ 2-3 week	60	manual	1 value/epoch

5.3.1 Grassland site (P360)

P360 (Figure 40) is one of 1100 GPS stations that make up the EarthScope Plate Boundary Observatory. It was installed with the purpose of studying crustal deformation in the western U.S. The typical setup is made of a 2-m tall metal tripod drilled into bedrock. At the apex rests a

choke-ring antenna with an L-band element in its center (boresight facing zenith), both housed within a radome.

P360 is located in an open field (Figure 41); visibility to the ground is unobstructed. The nearest trees are ~ 200 m away due west (Figure 42); visibility to the sky is also excellent. Thus reception of both reflected and direct signals is guaranteed, as required in this technique (Larson and Nievinski 2013). The ground is mostly flat, with topography deviating no more than 3.5 m above and below the mean horizontal plane within a 100-m radius of the antenna. Slopes do not exceed 5 degrees (at the 10-meter spatial scale), and aspect angle is such that the ground is facing N-NE. At the meter-level scale the terrain is rugged, with exposed rocks and littered with loose cobbles. Land cover classification is grasslands. There is a watercourse 200-m away due NE-W.

At P360 we show three years of estimated snow depth. Collection started at the time when L2C tracking was enabled in the receiver. Throughout this period there are sonic snow depth measurements available from a 12 ± 4 km distant 60-m higher-altitude SNOTEL station. In the third year there is up to four-times daily *in situ* validation data. Figure 45 shows results separately for each water-year (the period starting October 1st through September 30th of the following year, encompassing the northern-hemisphere winter; a water-year is designated by the calendar year in which it ends). For the second year, satellite PRN 25 had been launched, and for the third year, satellite PRN 01 was also launched; each new satellite adds potentially four new track clusters, as indicated in the year-to-year diagrams of clusters. This is not to say that all are equally good; the diagrams of track clusters shown in Figure 44include circles that are proportional to the weight of each cluster (i.e., larger circles indicate more important clusters)

The final GPS estimates of snow depth follow from an averaging of all available tracks, whose individual snow depth values were previously estimated independently. A new average is

produced twice daily utilizing the surrounding 1-2 days of data (depending on the data density), i.e., 12-h posting spacing and 24 to 48-h moving window width. The averaging interval must be an integer number of days, so as to minimize the possibility of snow depth artifacts caused by variations in the observation geometry, i.e., the number and azimuth of tracks being combined, which repeats daily.

The GPS estimate of the snow depth is shown as a thick solid red line. The dark-gray band denotes the 95% confidence interval for the average, which follows from the formal least-squares uncertainty, scaled by the RMS of residuals, then expanded based on Student's t-distribution for the number of tracks available. The light-gray band denotes the 95% simultaneous prediction interval for a random observation, utilized to detect and reject outliers. The relationship between the two statistics is analogous to the standard error of the mean $\sigma/\sqrt{n}$ and the standard deviation σ in an n-element sample. Individual GPS tracks that passed quality control are shown as gray dots.

SNOTEL sonic depth values are shown as a dashed blue line. Generally the timing of snowfall events is comparable across the ~ 10-km apart locations, although snow depth is not exact amount of snow is not. A salient type of temporal feature that is well captured by both GPS and SNOTEL is the sharp transition between accumulation and settling that happens when precipitation stops and snow depth starts to drop.

For the water-year 2012 there were *in situ* measurements, corresponding to a marked pole co-located with the GPS antenna, which was photographed up to four times daily. These are shown in green in the figure. We find a markedly improved agreement in terms of absolute amounts of snow depth, in contrast to the SNOTEL comparison. This type of *in situ* data is not expected to be exactly comparable with the GPS estimates. The pole data stem from one-time

readings at a fixed location with no coverage area, it lacks statistical replication necessary to quantify the variability of snow depth. Clearly, the pole readings are to be treated as individual observations rather than estimates of the mean. As such they are not expected to fall within the GPS confidence bounds for the mean. Furthermore, its block support or spatial footprint is orders of magnitude smaller than the GPS, so the GPS prediction bounds also do not apply. If the pole were displaced over multiple locations so as to realize a spatial average, its confidence interval could become as small as to that of the GPS, but its prediction interval would remain wider than that of the GPS, as the pole is sensitive to smaller-scale snow depth variations that are smoothed out in the GPS sample. With a single pole at a fixed location, we rely on temporal variability to produce the equivalent of ensemble averaging (ergotic assumption). This is found to generally improve the comparison, with the main exception happening between days 100 and 110, during which the GPS estimates are particularly stable, suggesting the quiet temporal dynamics fail to provide adequate randomization.

During the snow-free period we find that reflector height is not exactly zero. Variations occur mainly when the scattering medium is transitioning, from snow to slush/mud and eventually grass-covered soil. This issue is both a challenge and an opportunity. On the one hand, it poses the risk of being mistaken for snow depth events. In fact, the identification of the site-overall (i.e., non-cluster specific) zero-level or bare-ground reflector height is perhaps the weakest link in the whole GPS processing chain of snow depth retrieval, as it relies on only a few data points (we compute it as the 5^{th} percentile of average reflector heights over the snow-free period). On the other hand, such observations attest to the prospects of using GPS reflector heights for monitoring environmental targets other than snow, such as vegetation biomass. If

successful, the estimation of non-snow targets would contribute to guaranteeing that snow depth remains non-negative.

Figure 46 highlights the inter-annual variations in snow depth at the P360 site. Although the snow season begins consistently in the second half of November, the end of the season is markedly different, deviating by 15-30 days from year to year.

5.3.2 Forested site (RN86)

GPS site RN86 (Figure 47) was installed at the T. W. Daniel Experimental Forest, an area administered by Utah State University and the USDA Forest Service. The goal was to assess the impact of partial obstruction by the surrounding trees (Figure 48 and Figure 49) on snow depth retrievals. Ground crews manually collected *in situ* measurements around the GPS antenna approximately every other week. Measurements are made every 1-2 m from the GPS up to 25-30 m at six azimuths (0, 45, 135, 180, 225, and 315 degrees). Therefore it is possible to obtain *in situ* average estimates, with their own uncertainties (based on the number of measurements), which allows a more meaningful comparison. The first four as well as the last data collection visits were surveys of opportunity that did not follow this standard sampling protocol.

The diagram of track clusters (Figure 51) indicates a reduced visibility at the current site, compared to the previous open-field site. Also their uncertainty is not as favorable, as indicated by the small-radius circles. Furthermore, clusters are concentrated due south, with only two clusters located within ± 90° of north. Therefore, the GPS average snow depth is not necessarily representative of the azimuthally symmetric component of the snow depth. In the presence of an azimuthal asymmetry in the snow distribution around the antenna, the GPS average is expected to be biased towards the environmental conditions prevalent in the southern quadrant. In any case the comparison shows generally excellent agreement between GPS and *in situ* data (Figure

52). The first four and the last one *in situ* data points were collected with coarser spacing and/or smaller azimuthal coverage, which may be partially responsible for different performance in the first and second halves of the snow season.

The correlation between GPS and *in situ* snow depth at RN86 amounts to 0.984, indicating a very strong linear relationship. Inspecting the individual differences at each of 9 visits (Table 3), we find that all but one is within the corresponding uncertainty (obtained via propagation of the GPS and the *in situ* uncertainties). Carrying out a regression, the intercept 13.8 ± 13.7 cm is just barely statistically significant vis-à-vis its 95% confidence interval, although the latter is practically as large as the former. The regression slope 0.818 ± 0.132 m/m is significant by a wider margin: the expected value is ~ 6 times greater than its uncertainty. This non-zero slope could indicate a genuine systematic under-estimation on the part of the GPS compared to *in situ*, although there is the possibility of an artifact caused by different sampling coverage. A more detailed analysis, in which azimuthal variations are resolved, is warranted. The RMS of snow depth residuals improves from 12.2 cm to 6.8 cm after the regression. The finding about GPS under-estimating in situ snow depth should not extrapolated to other settings until the possibility of an azimuthal asymmetry artifact is ruled out.

The SNOTEL sensors are exceptionally close to the GPS at this site, only ~ 350 m horizontal distance and with negligible vertical separation. Yet the former is located under forest cover, while the latter is located in the periphery of the forest and senses the reflections scattered off the adjacent open field. Therefore only the timing of snowfall events agrees well, not the amount of snow. Although forest density is generally negatively correlated with snow depth (Gray and Male 1981), exceptions are not uncommon (López-Moreno and Latron 2008; Veatch

et al. 2009), especially in localized clearings exposed to intense solar radiation, where shading of the snow by the trees reduces ablation.

Table 3: GPS and *in situ* snow depth differences at the forested site (RN86) versus day of water year (DOY-W). Values are in centimeters; 95% uncertainty intervals are denoted as "±".

DOY-W	Difference	±	GPS	±	*In situ*	±
49	9	11	36	4	27	11
97	13	17	64	3	51	17
134	-23	13	100	10	123	9
148	-6	11	134	5	140	10
168	-10	12	124	4	135	11
183	-4	9	119	2	123	9
195	-6	12	88	4	94	11
209	-5	12	60	3	65	11
224	18	19	23	8	5	17

5.3.3 Alpine site (NWOT)

We now consider the data collected at the Niwot Ridge Long-term Ecological Research site in Colorado (Figure 53). At 3,500-m altitude, it is located in a saddle-like mountain-top, in an alpine tundra environment (Figure 54). A 3-meter tall continuously-operating GPS system was established in 2009 (Figure 55). Poles are staked at 50-m intervals making up a 120-by-400 meter Cartesian grid at which snow depth is measured manually using a snow sampling tube, approximately every two weeks; we use *in situ* data collected at the pole nearest to the GPS (shown in Figure 53). The ground at the present site is not as planar as in the previous two sites, but visibility to the sky is good, with no trees and only minor topographical obstructions, predominantly due east and west. Indeed, the diagram in Figure 57 shows a nearly uniform azimuthal distribution of track clusters.

The nearest SNOTEL location is more than 4 km away and ~ 380 m lower in altitude. As is typical for SNOTEL stations, this one is found among trees, whereas the GPS is above the tree line. Therefore their comparison comes with caveats. Similar as for the forested site, the

comparison against co-located *in situ* is more favorable. Results for the pole co-located with the GPS antenna is good agreement for the first two years, less so in the third year (Figure 58).

In the last year the peak snow depth is much smaller than in previous years, and the performance of the GPS deteriorates. This is partially because the amount of snow is not sufficient so as to fill in the ground depressions during most of the season. Therefore the multiple GPS track clusters are not as comparable as in previous years, when the inter-cluster variability leveled out as the air/snow interface became more planar than the snow/ground interface. Furthermore, the spatial variability affects the *in situ* data, too, as the single pole co-located with the GPS is not as representative of the area sampled by the GPS. The remaining poles are more than 50 m from the GPS, so their potential contribution is questionable under such low-snow conditions.

The inter-annual variations in snow depth at the NWOT site (Figure 59) are more drastic than at P360. Indeed, it exhibits a five-fold difference in peak snow depth, from ~ 0.5 m in 2011-2012 up to ~ 2.5 m in 2010-2011. The timing of the end of the season varies by more than a month over this three-year period. The exact beginning of the season is less clear as the snow that accumulates from the initial precipitation events can be totally dissipated if the snowpack is not replenished with more frequent and vigorous snowfalls.

Figure 60 shows a scatterplot of the GPS vs. *in situ* snow depth for the entire three-year period at NWOT. The correlation is 0.980, which indicates a very strong linear relationship. Carrying out a regression, the intercept 2.1 ± 2.8 cm is not statistically significant vis-à-vis its 95% confidence interval. The regression slope 0.89 ± 0.03 m/m is significant, which means that at this site the GPS estimates are lower than *in situ* snow depth by about 10%, although their footprints are not overlapping.

The RMS of residuals improves from 10.7 +/- 3.0 cm to 7.8 +/- 2.6 cm after the regression; the RMS confidence interval reflects the degree of freedom (60 observations minus 2 parameters) and the sample kurtosis. Both pre- and post-fit residuals are skewed to the right, i.e., their histogram (not shown) have a longer positive tail, typical of data having a lower bound of zero, as is the case for non-negative snow depth. In fact, residuals are more dispersed at smaller snow depth values, especially when considered in proportion or relative to the *in situ* values (Figure 60, bottom panel).

5.4 CONCLUSIONS

We have demonstrated a statistical inverse model for estimating snow depth based on GPS multipath employing SNR measurements collected at three sites over a multi-year period. The model performance was assessed against independent *in situ* measurements and found to validate the GPS estimates, to within the limitations of both GPS and *in situ* errors. We now distill the lessons learned in developing and applying the model.

The suitability of GPS SNR measurements for snow monitoring was found to be heavily influenced by the site conditions; this lends weight to the finding of Larson and Nievinski (2013) that clearance to the satellite line-of-sight as well as to the ground are strong requirements for GPS multipath reflectometry (GPS-MR). Therefore the quality of retrievals may vary enormously over different azimuths at the same site. Furthermore, for a high-quality track cluster, a few percent of its daily retrievals should be discarded for a variety of reasons, such as receiver failures and heavy snowfall. Quality control (QC) is therefore mandatory for operational exploitation of GPS-MR.

A multi-test QC strategy – including goodness-of-fit, reflector height uncertainty, peak elevation angle, and statistical degree of freedom – was found to work best, as no single test

detected all outliers. Combining multiple track clusters further improved the precision of GPS retrievals, yielding daily averages that captured remarkably well the temporal dynamics of snow accumulation and ablation, including sudden changes associated with new snow. Inter-cluster combination also provided a second opportunity for QC, complementing the previous intra-cluster QC. Statistically robust methods – e.g., median instead of the mean – were adequate in achieving a reasonable level of processing automation and dispense with frequent manual intervention. Continuity of the time series as new satellites were launched every year indicate no obvious satellite-dependent biases; this stability is paramount for future utilization of GPS-MR results in climate studies.

Turning attention to aspects that would require more care in the future, our treatment of the azimuthal asymmetry exhibited by snow depth was admittedly cursory, in the sense that we only tried to minimize its impact on the daily site averages by making clusters more comparable. This treatment worked well when the amount of snow was enough to fill in the ground depressions. Yet when the amount of snow was insufficient to make the air/snow surface more planar than the snow/ground surface (alpine site, 2011-2012), the treatment failed to improve the dispersion around the site average.

Another aspect that would deserve closer scrutiny is the statistics of outliers. As their density increases, the performance of the estimator degrades, eventually reaching a breakdown point when outliers no longer form the minority of the data. The outlier rejection could be more aggressive, particularly at the forested site (RN86), although this needs to be balanced against excessive and unnecessary gaps in the time series.

Finally, a delicate matter that should be further investigated is the definition of the zero-depth line or the average ground height. The challenge is that although we can measure reflector

heights precisely, we cannot unambiguously attribute an individual track estimate or even a daily site average to a specific target, i.e., to distinguish between snow vs. vegetation vs. soil moisture changes manifesting in reflector height. We have relied on the temporal dynamics of reflector heights along with reasonable assumptions about the snow behavior and optionally ancillary information (photographs, temperature records, climatic expectations, etc.) to determine the snow-covered period. This strategy worked very well for large amounts of snow, but it becomes less reliable for smaller amounts. As a rule-of-thumb, 10-cm reflector height change would be a reasonable cutoff value for distinguishing snow from other targets. So this issue is more serious for ephemeral snow sites, but it remains relevant for all sites, as GPS could be systematically under- or over-estimating snow depth. Keeping the performance in perspective, a 10-cm error would still be less than 5% for the 2.5-m snow depths observed at the alpine site in 2010-2011.

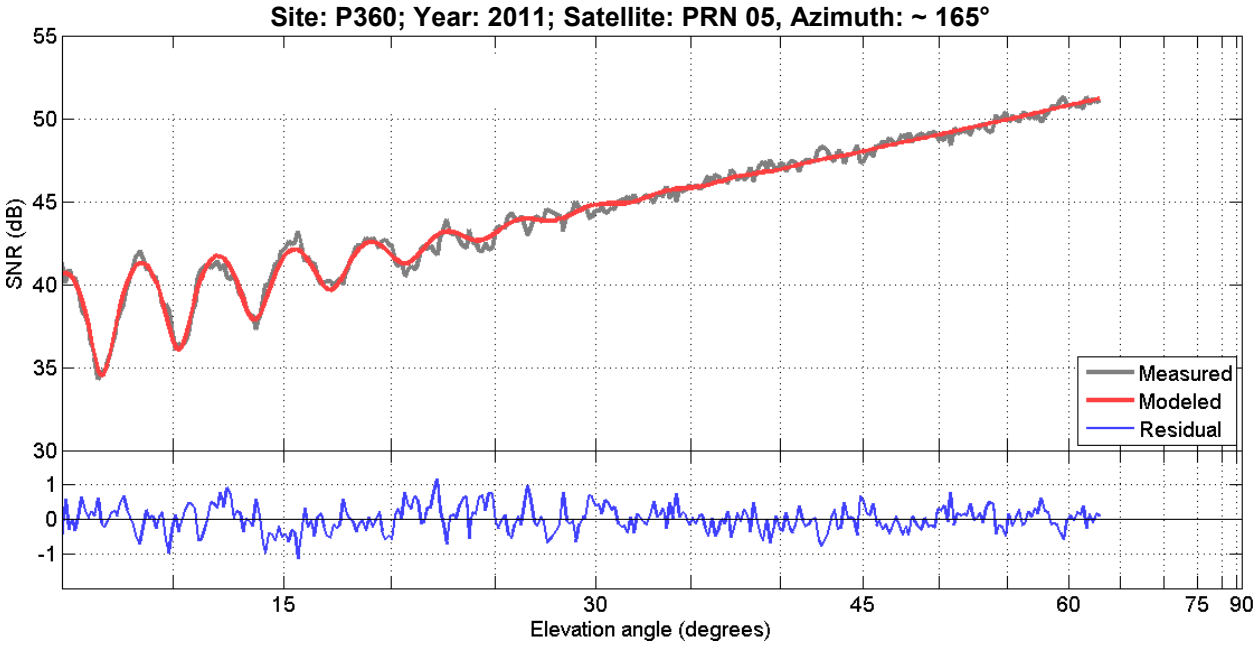

Figure 28: Example of good observation fit.

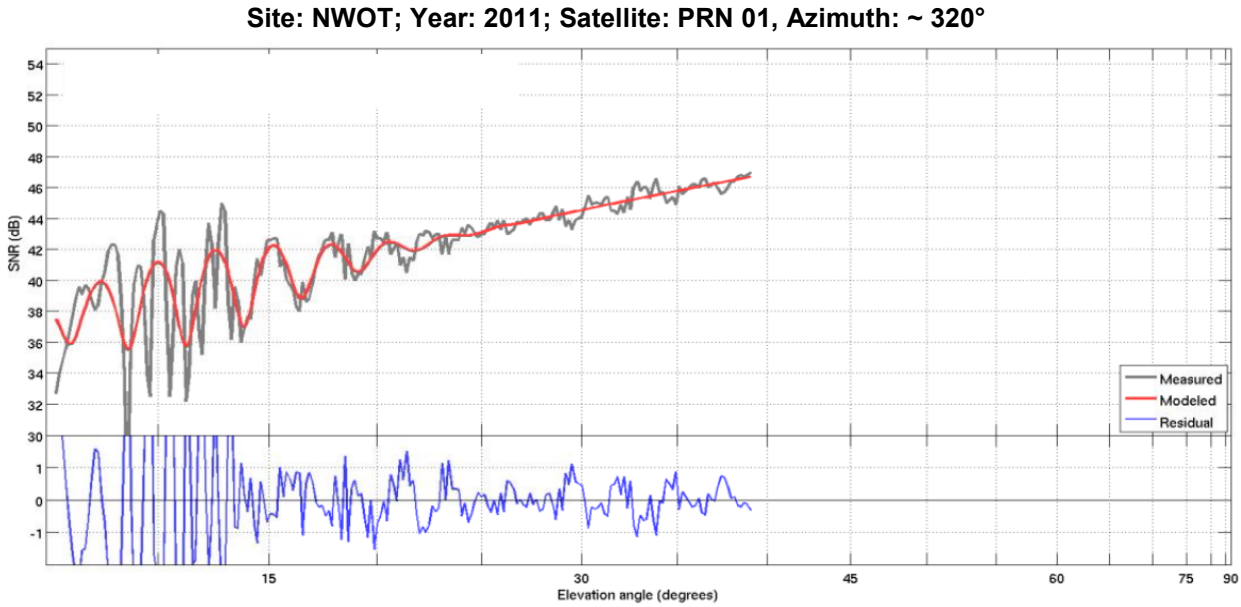

Figure 29: Example of secondary reflections.

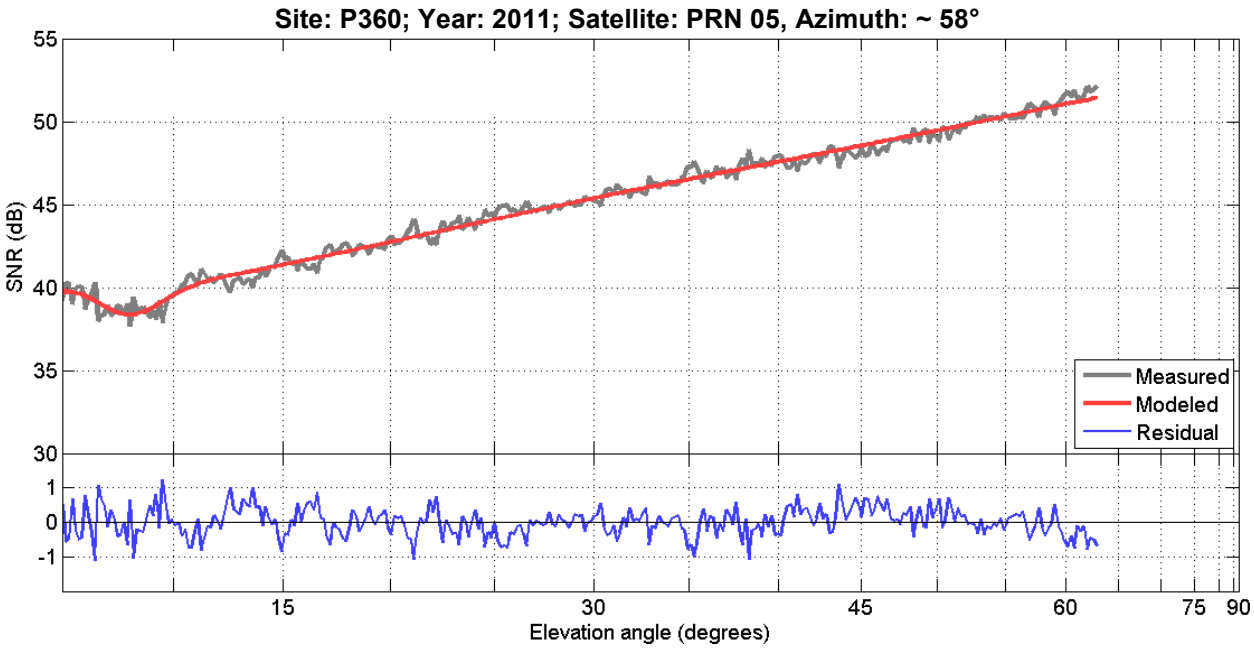

Figure 30: Example of vanishing interference fringes.

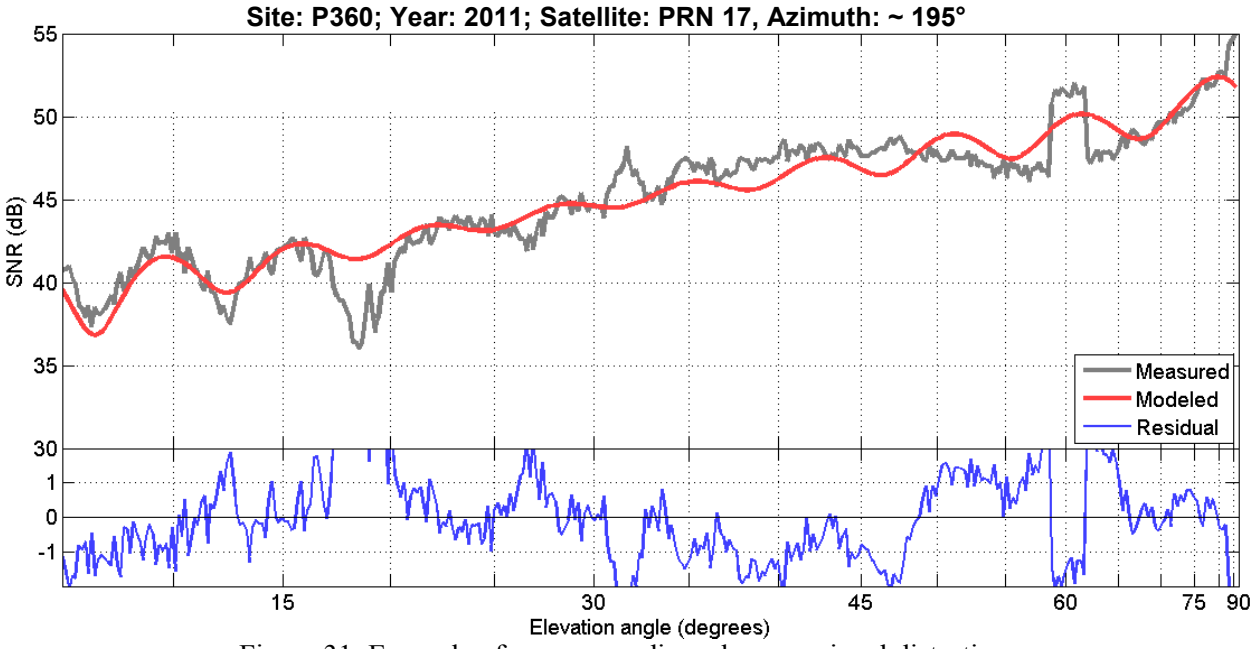

Figure 31: Example of more sporadic and severe signal distortion.

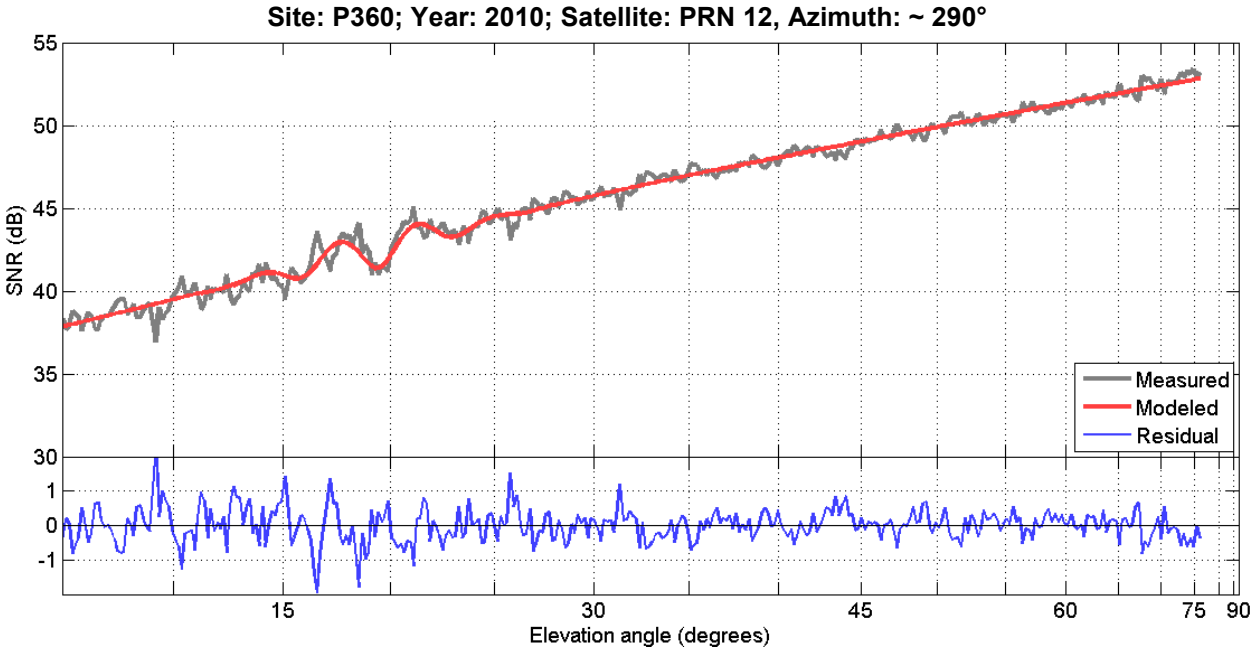

Figure 32: Example of non-typical interference fringes.

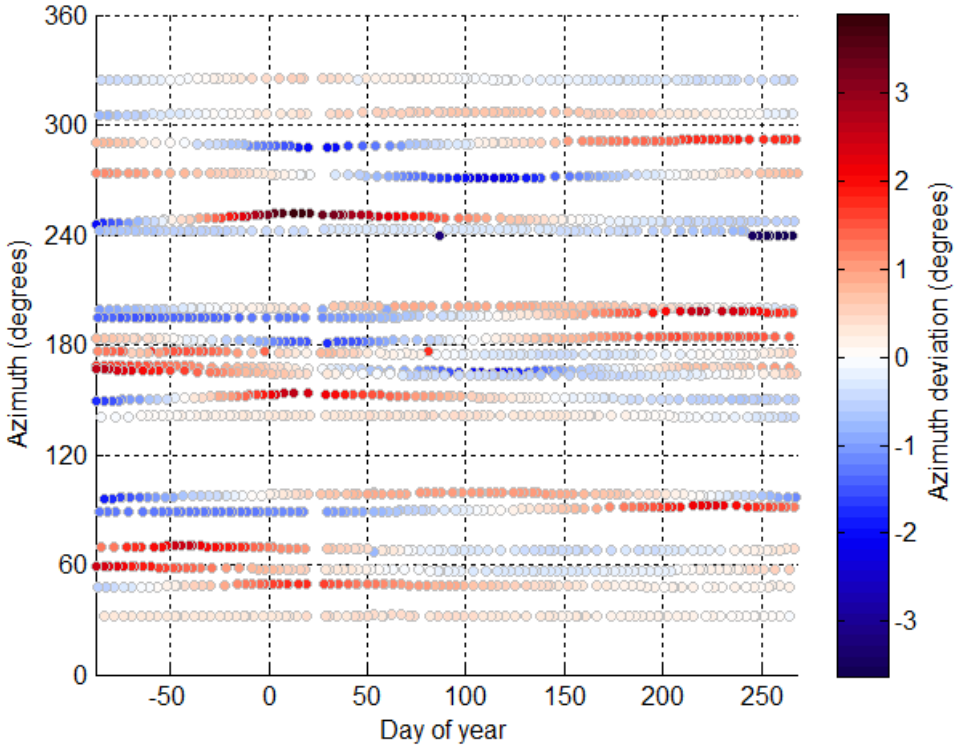

Figure 33: Satellite track azimuths at a fixed 10° elevation angle, as well as their deviation from the median, as observed at the grassland site (P360) in the water-year 2011.

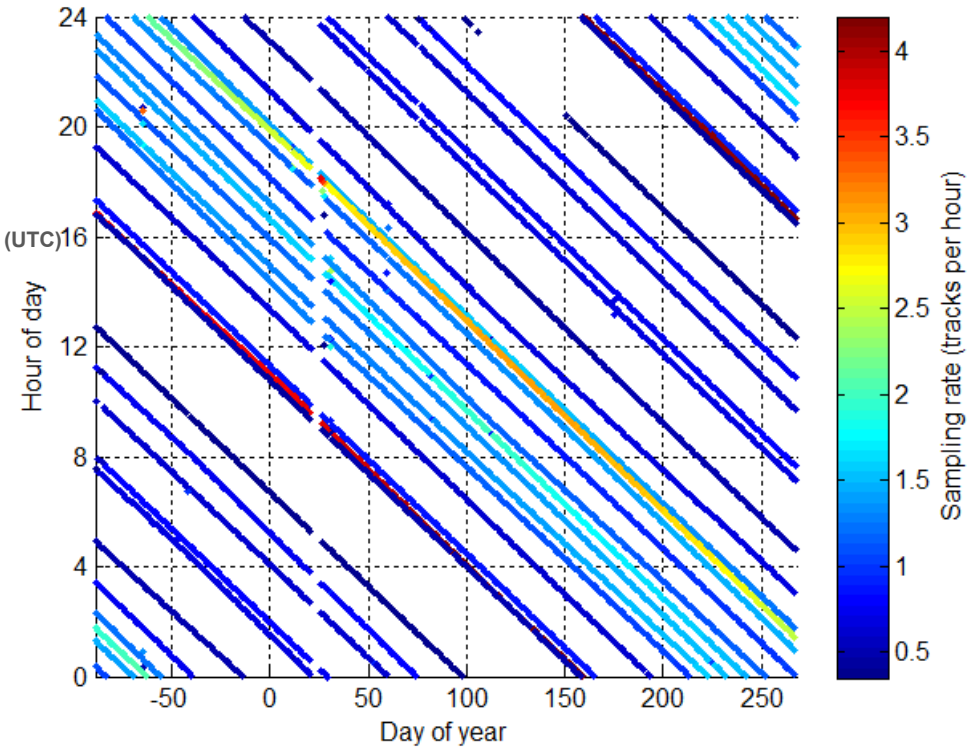

Figure 34: Satellite track hour of the day (UTC) over day of year, as observed at the grassland site (P360) in the water-year 2011.

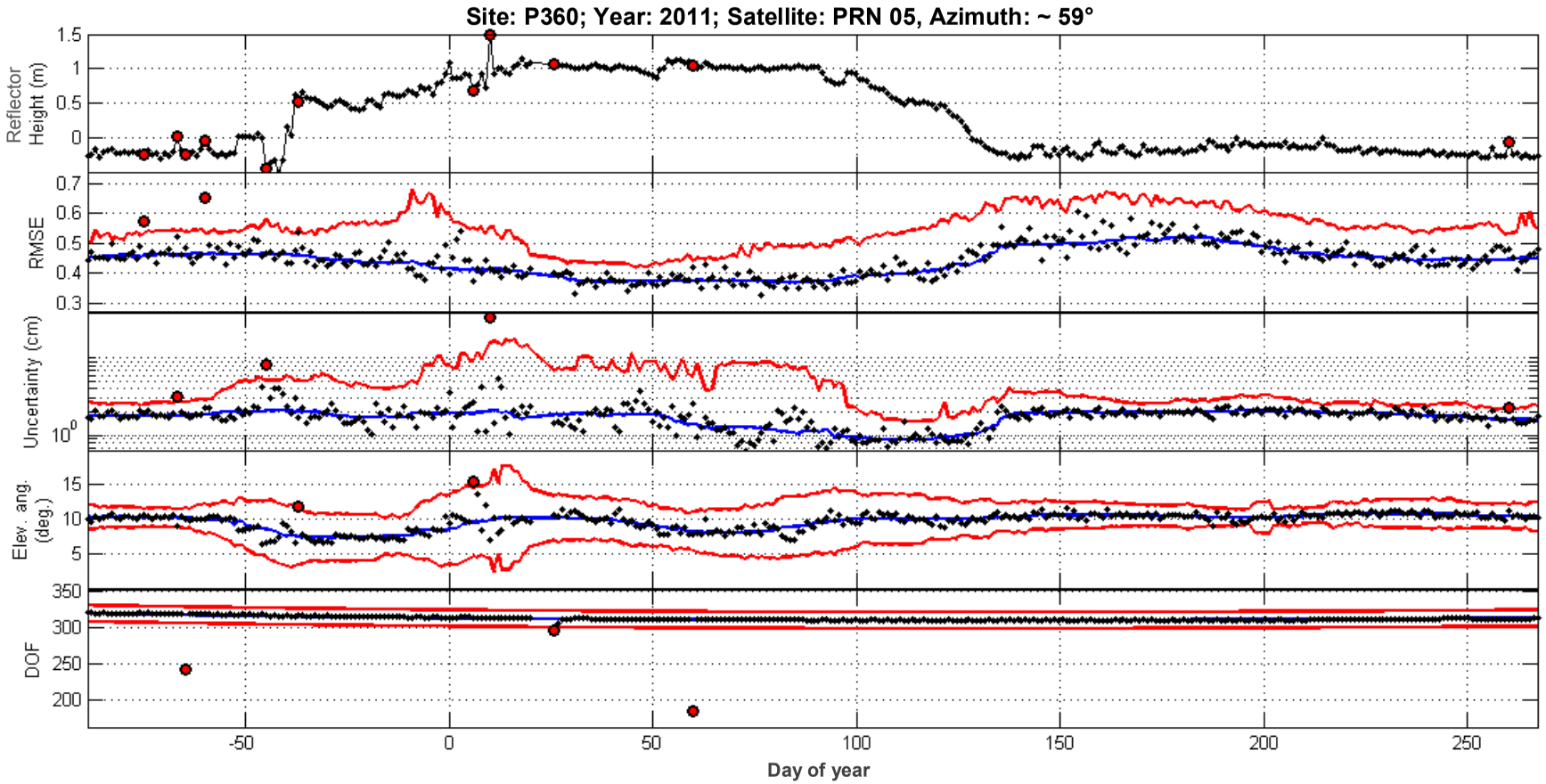

Figure 35: Time series of quality control tests for a single track cluster (satellite PRN 05, azimuth ~ 59°, ascending), as observed at the grasslands site (P360) in the water-year 2011. In each panel, black dots are independent day-to-day track retrieval statistics. Blue and red lines represent, respectively, the expected tendency and dispersion, both based on a 15-day moving average. The bounds are two-sided for normally-distributed variables (DOF and peak elevation angle) and one-sided for χ^2-distributed variables (uncertainty and RMSE). Red circles are retrievals deemed to be outliers, separately for each test in the lower four panels, and in conjunction for all tests, in the top panel

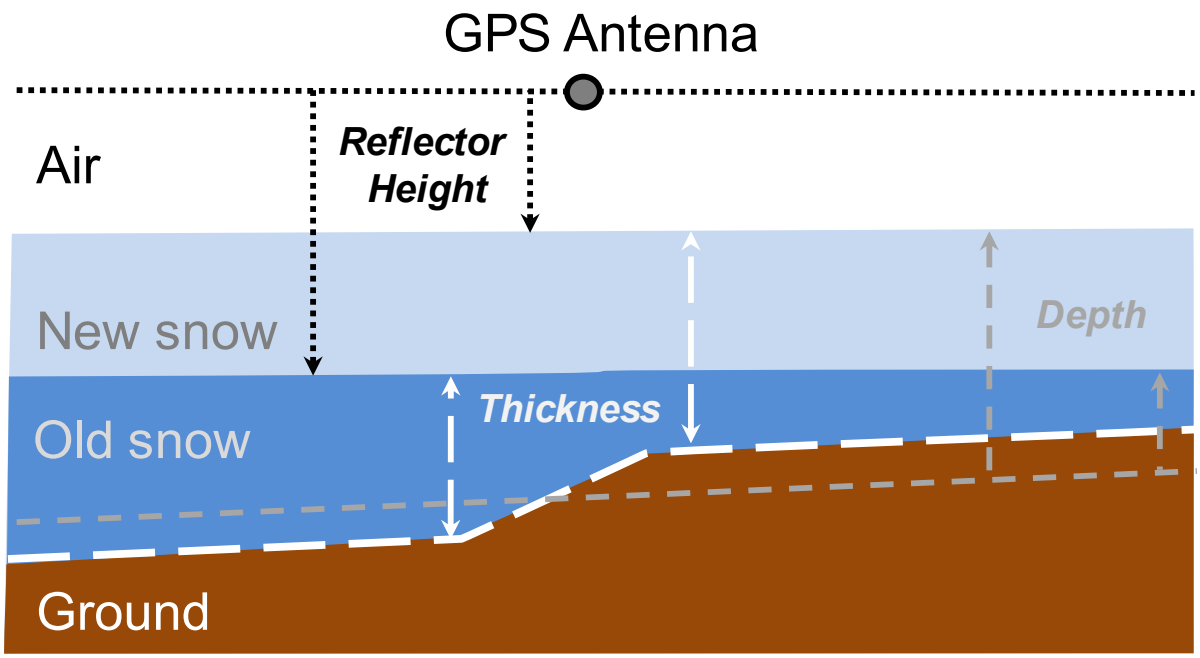

Figure 36: Snow reflector height, snow thickness, and snow depth as different vertical coordinates

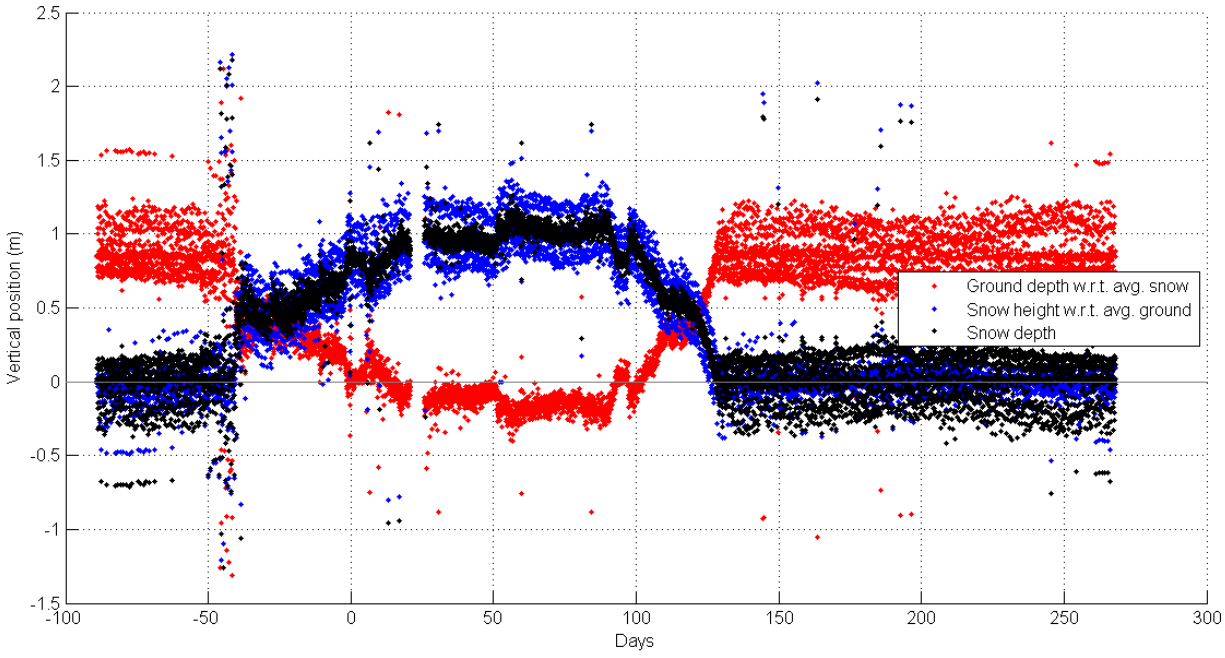

Figure 37: Raw snow retrievals – before quality control – expressed in terms of reflector height (in red), thickness (in blue), and depth (in black), as observed at the grassland site (P360) in the water-year 2011.

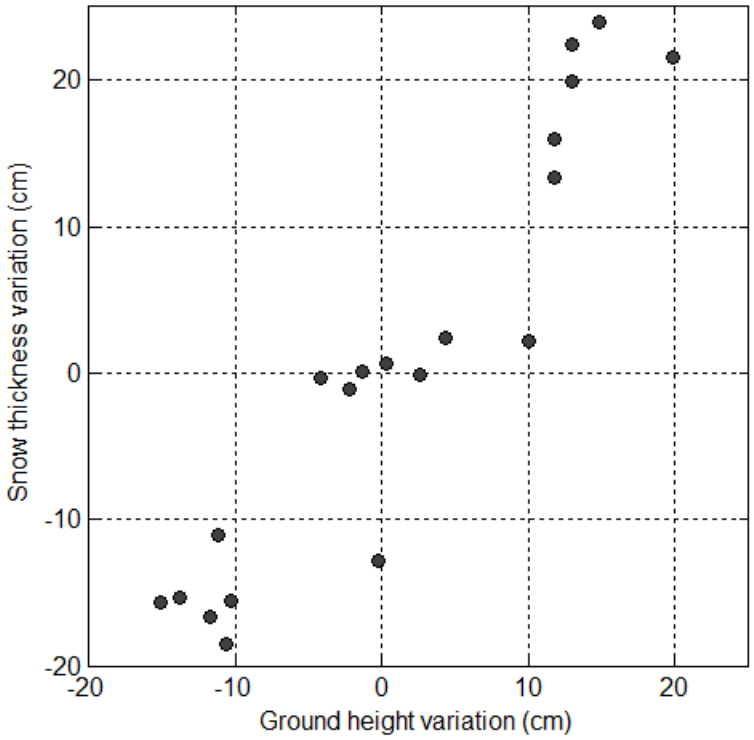

Figure 38: Correlation between inter-cluster variations in snow thickness and in ground height, as observed at the grassland site (P360) in the water-year 2011.

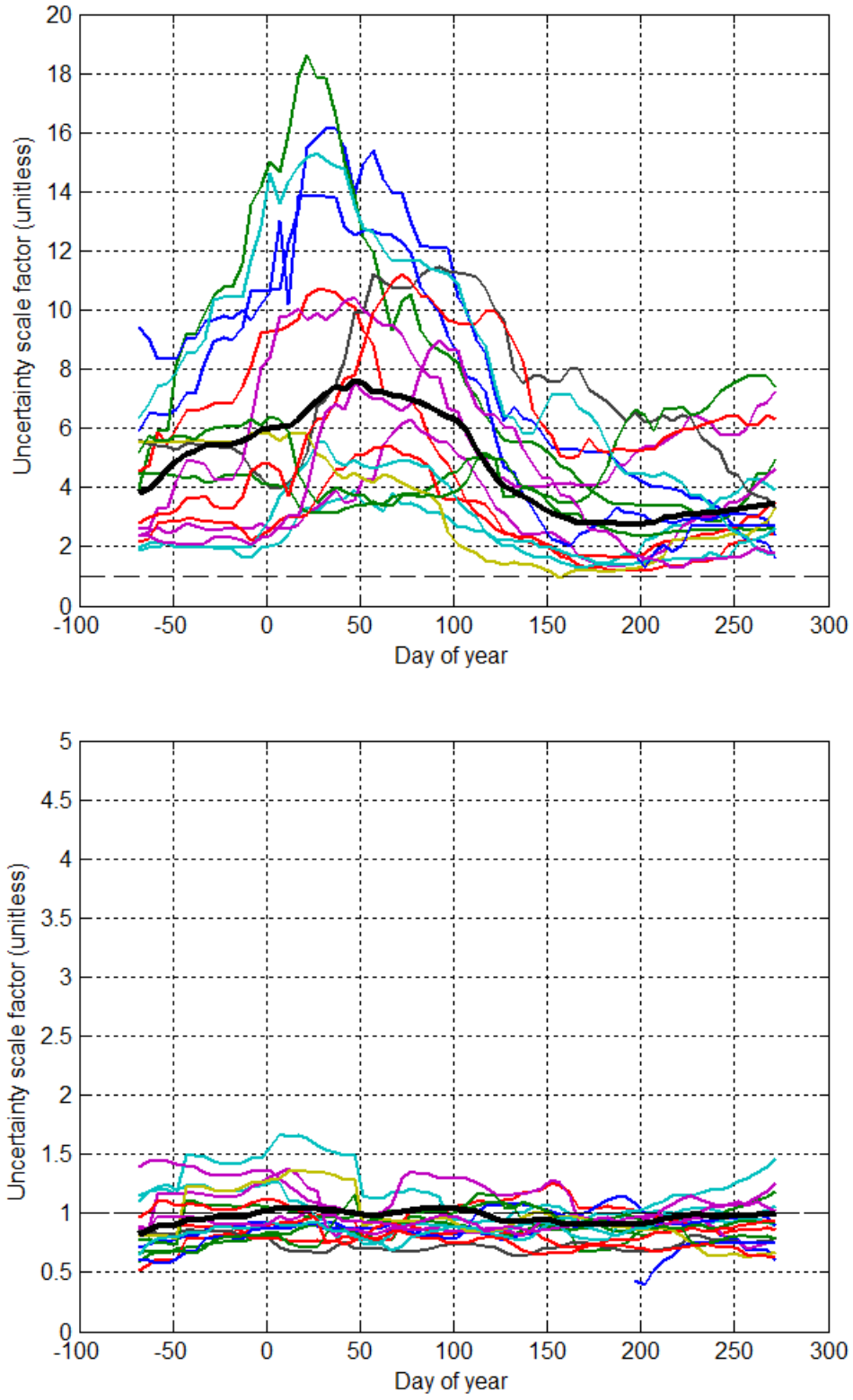

Figure 39: Uncertainty scaling factors, as observed at the grassland site (P360) in the water-year 2011; top: as a result of the first iteration; bottom: as a result of the second iteration. Each trace represents a unique track cluster.

Figure 40: Ground conditions in the vicinity of the GPS antenna at the grassland site (P360).

Figure 41: Sky visibility in the surroundings of the GPS antenna at the grassland site (P360).

Figure 42: Aerial view of the grassland site (P360) around the GPS antenna (marked with a circle).

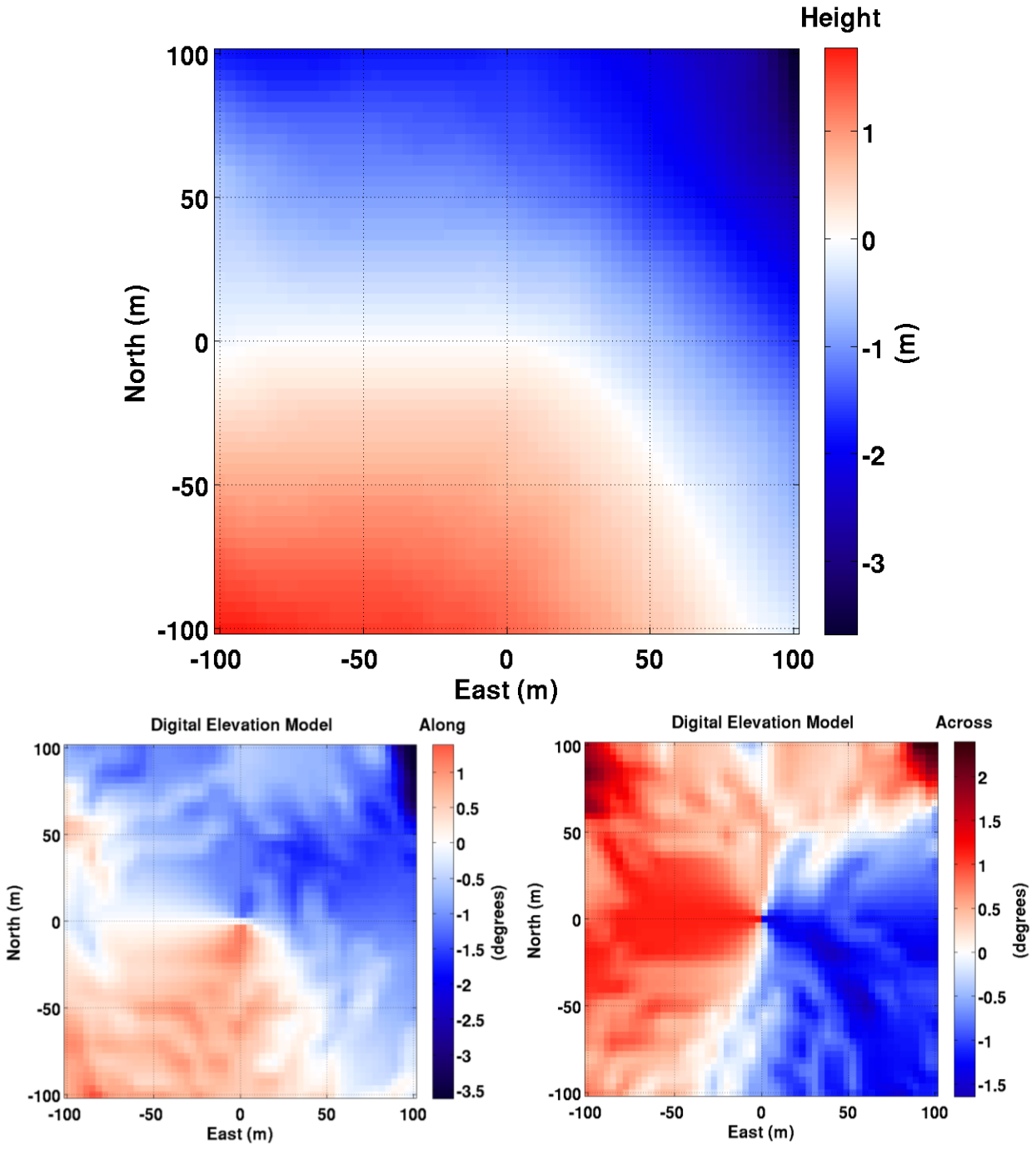

Figure 43: Ground at the grassland site (P360); surface height (top), along-track or radial tilting (bottom left), and across-track or azimuthal tilting (bottom right). The GPS is at the origin.

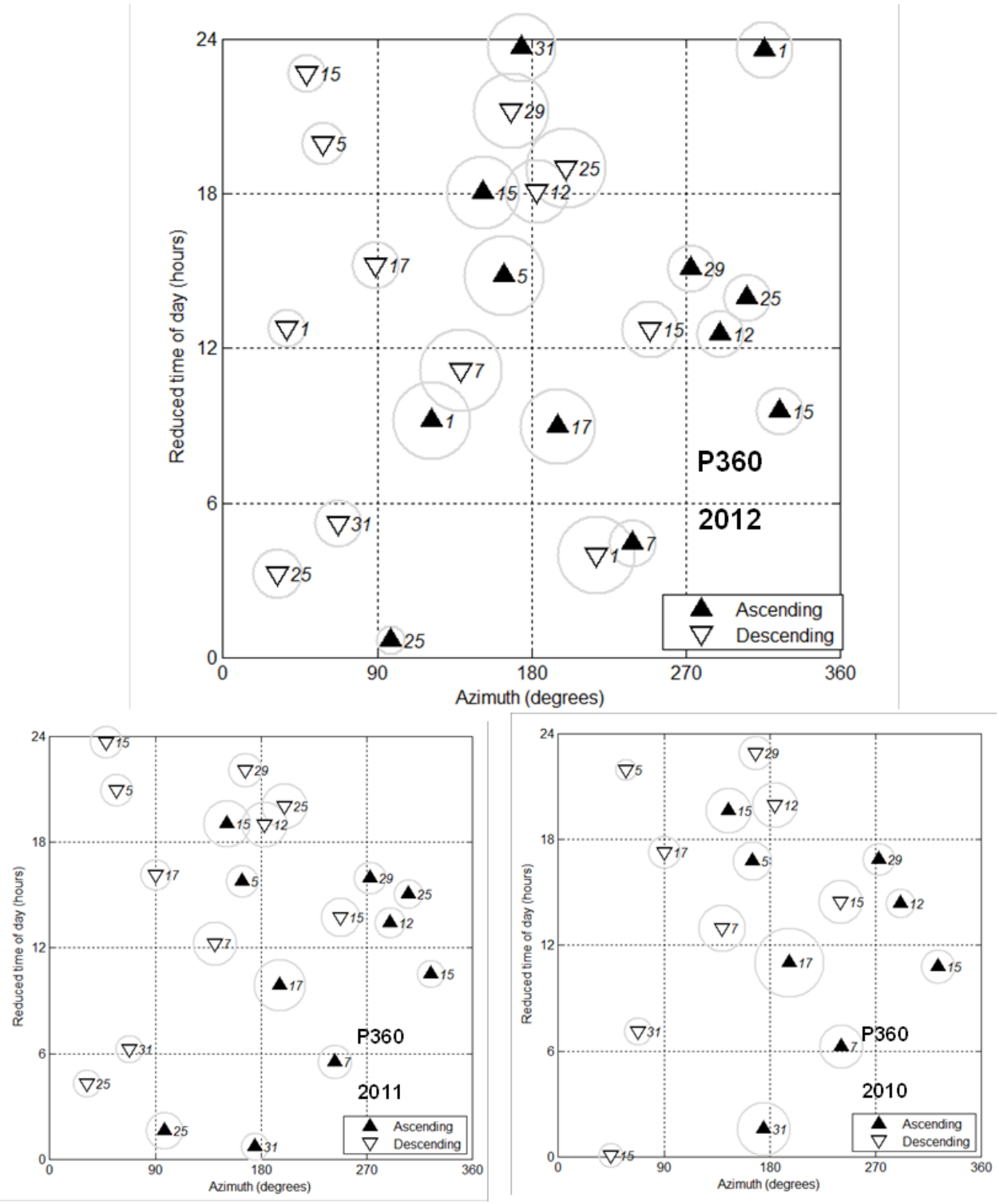

Figure 44: Diagram of satellite track clusters available at least 50% of the year at the grassland site (P360), for three water-years: 2012 (top), 2011 (bottom left), and 2010 (bottom right). Italic labels denote satellite PRN number.

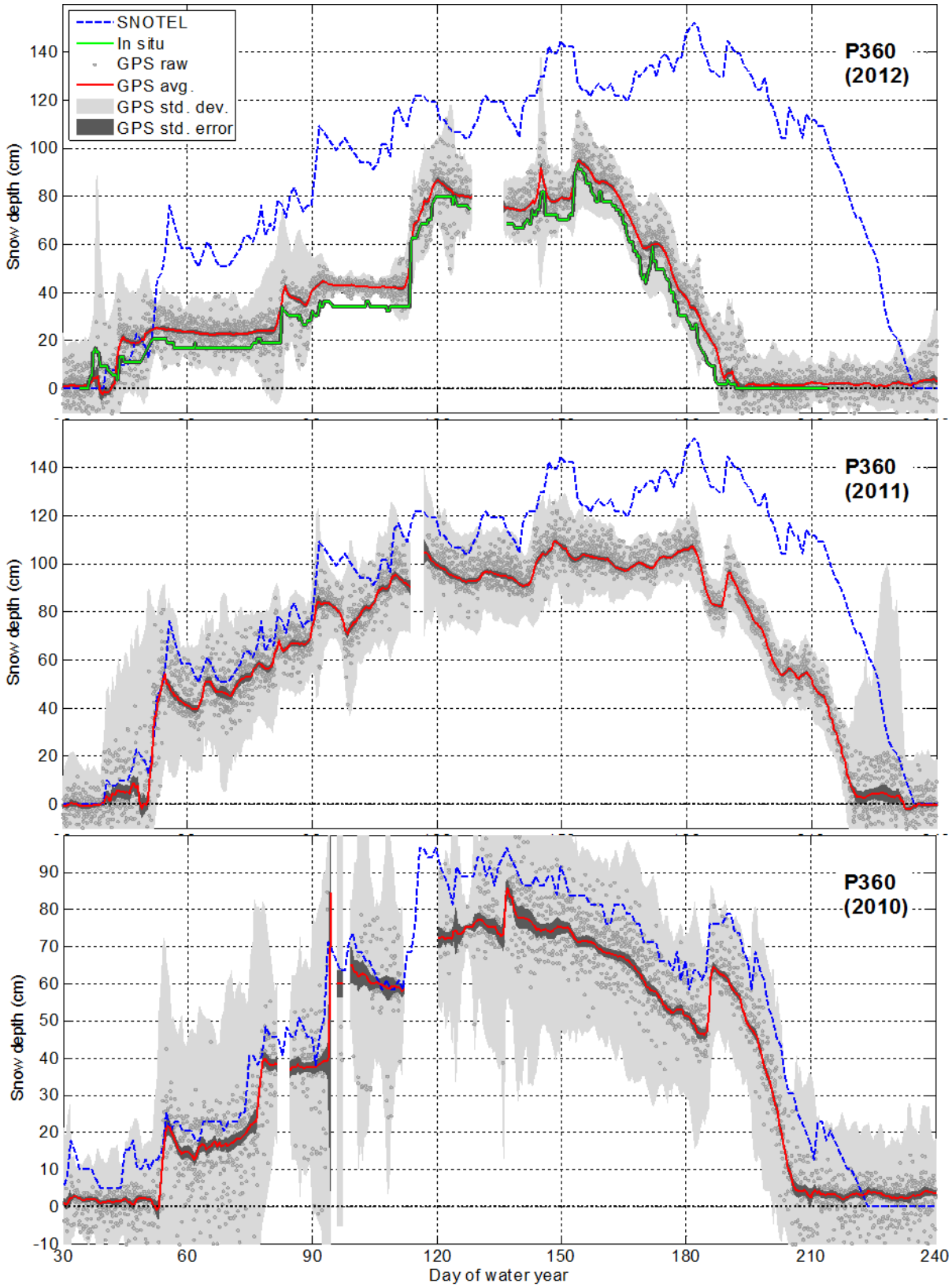

Figure 45: Snow depth measurement at the grassland site (P360) for three water-years.

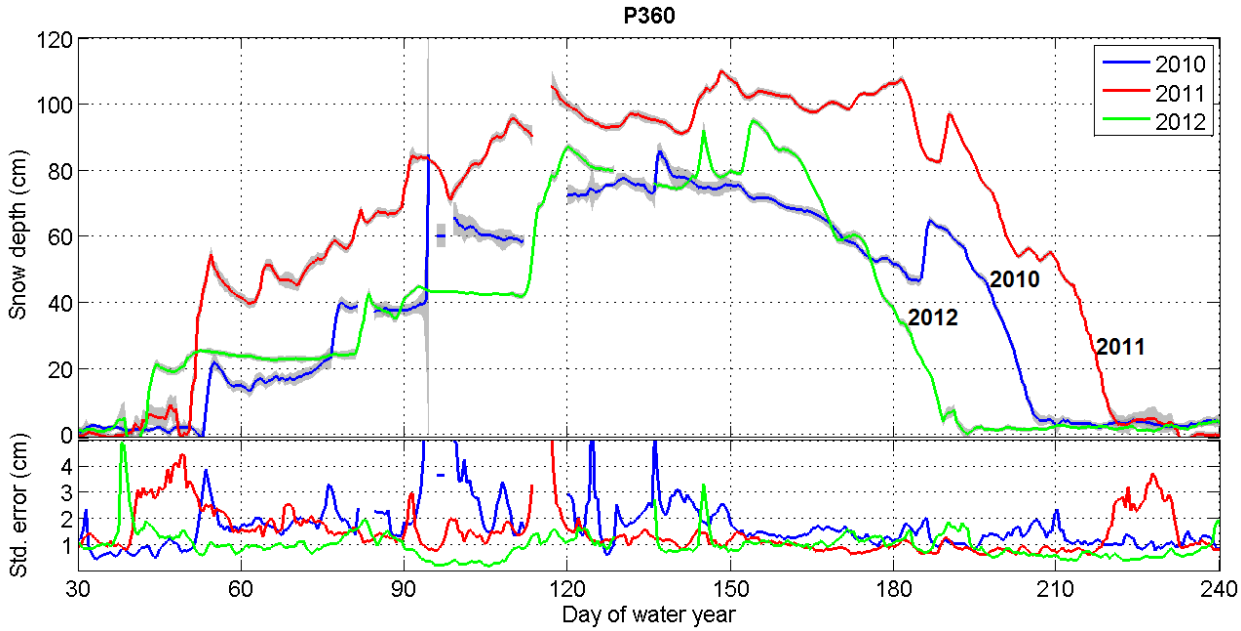

Figure 46: GPS estimates over multiple years at the grassland site (P360); the standard error of the mean is shown as gray bands in the top panel and also separately in the bottom panel.

Figure 47: Ground conditions in the vicinity of the GPS antenna at the forested site (RN86).

Figure 48: Sky visibility in the surroundings of the GPS antenna at the forested site (RN86), highlighting the proximity to trees in the north side; due south (not shown) visibility is less obstructed.

Figure 49: Aerial view of the forested site (RN86) around the GPS antenna (marked with a circle).

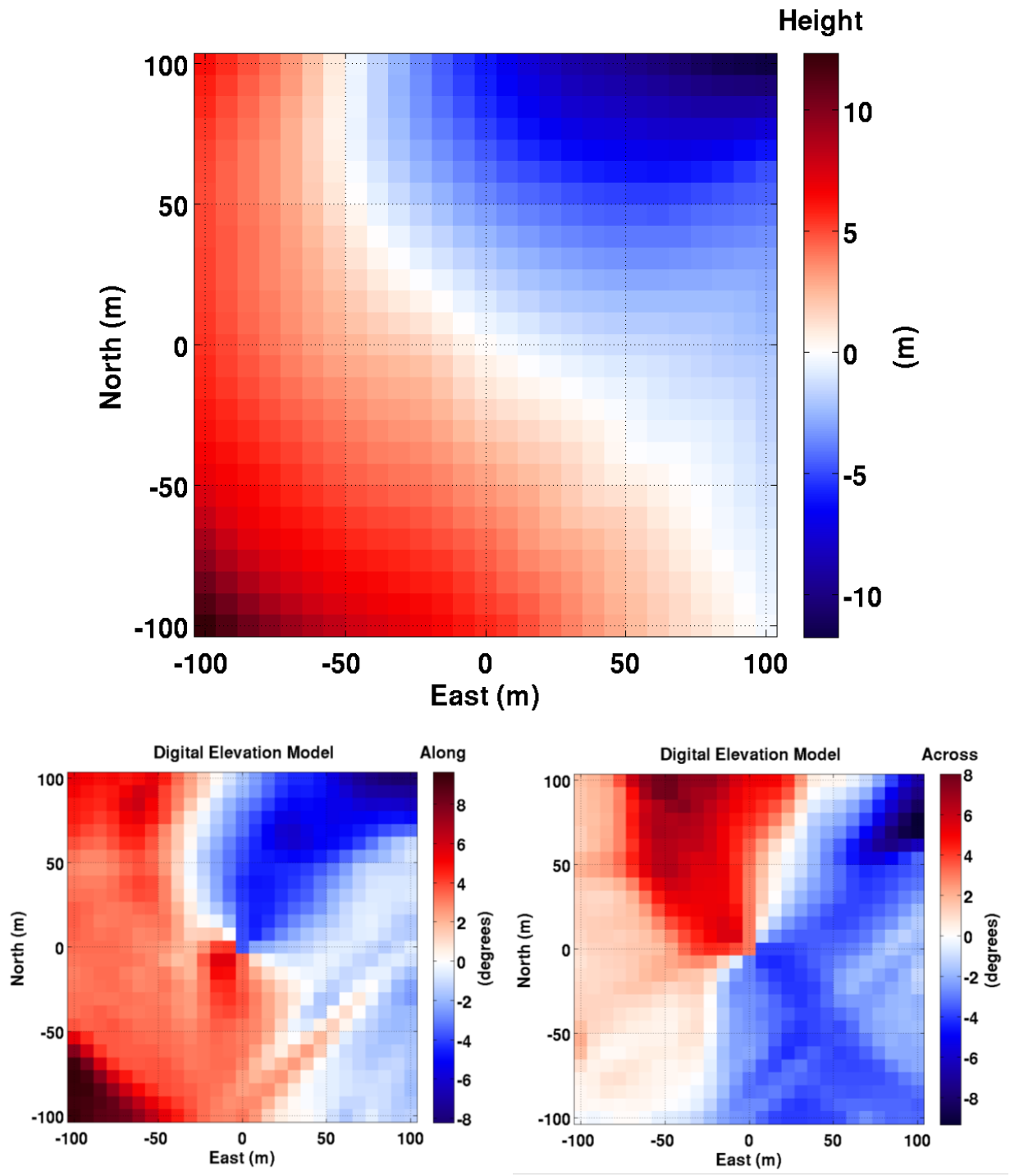

Figure 50: Ground at the forested site (RN86); surface height (top), along-track or radial tilting (bottom left), and across-track or azimuthal tilting (bottom right). The GPS is at the origin.

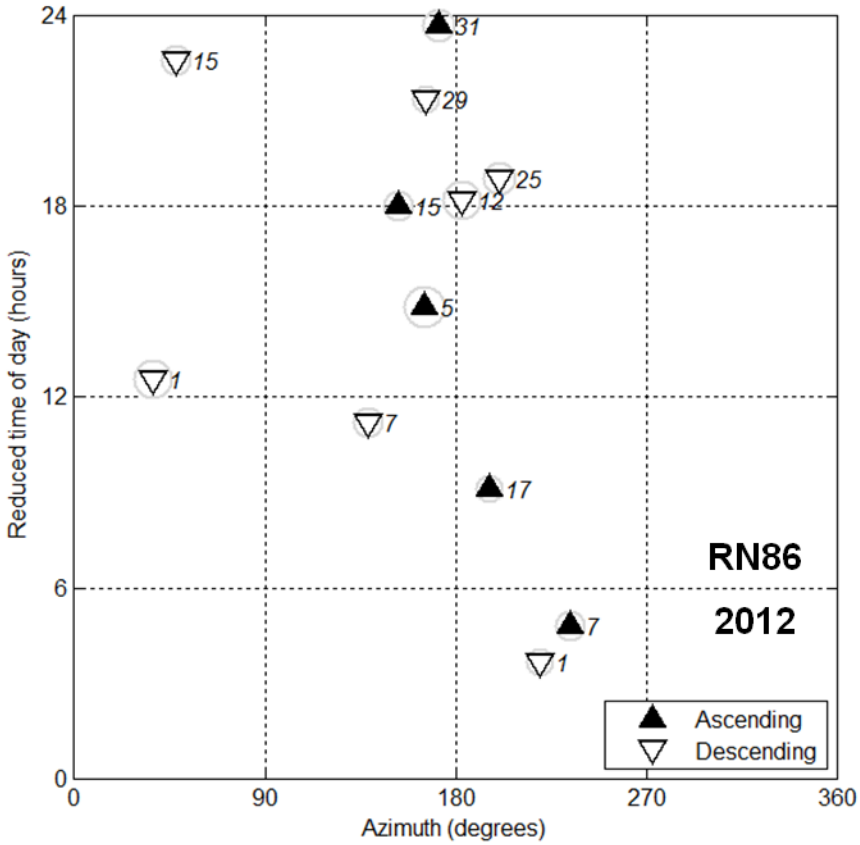

Figure 51: Diagram of satellite track clusters available at least 50% of the year at the forested site (RN86).

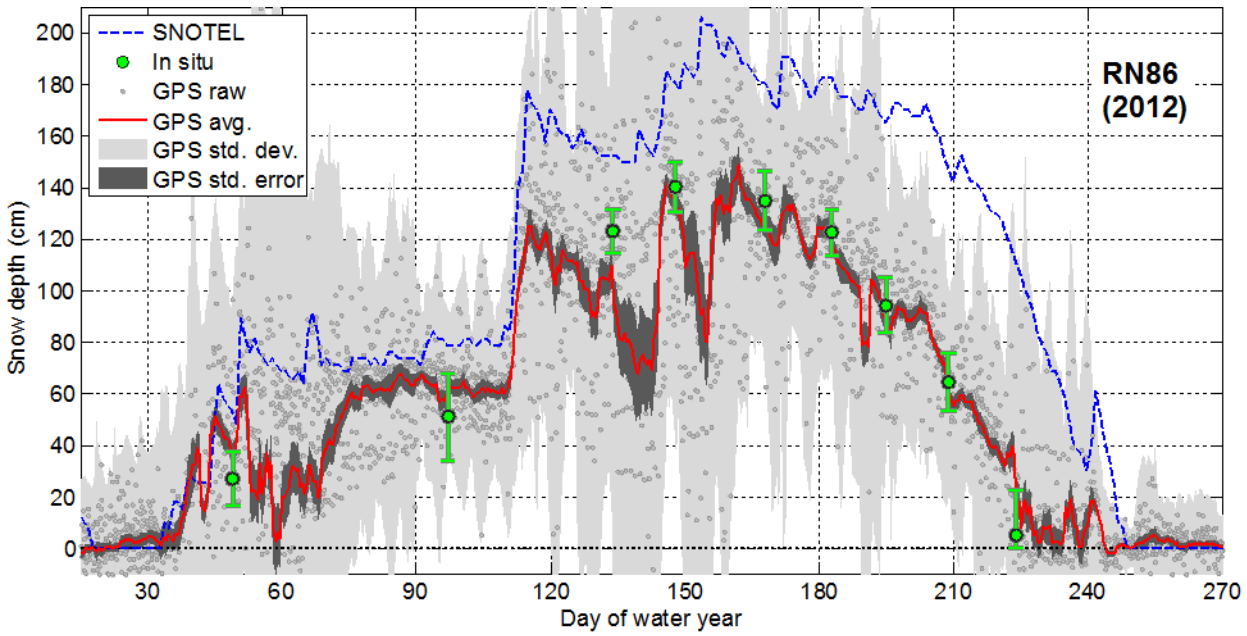

Figure 52: Snow depth measurement at the forested site (RN86) for the water-year 2012; see text for discussion and description of details.

Figure 53: Ground conditions in the vicinity of the GPS antenna at the alpine site (NWOT).

Figure 54: Sky visibility in the 360° surroundings of the GPS antenna at the alpine site (NWOT).

Figure 55: Aerial view of the alpine site (NWOT) around the GPS antenna (marked with a circle).

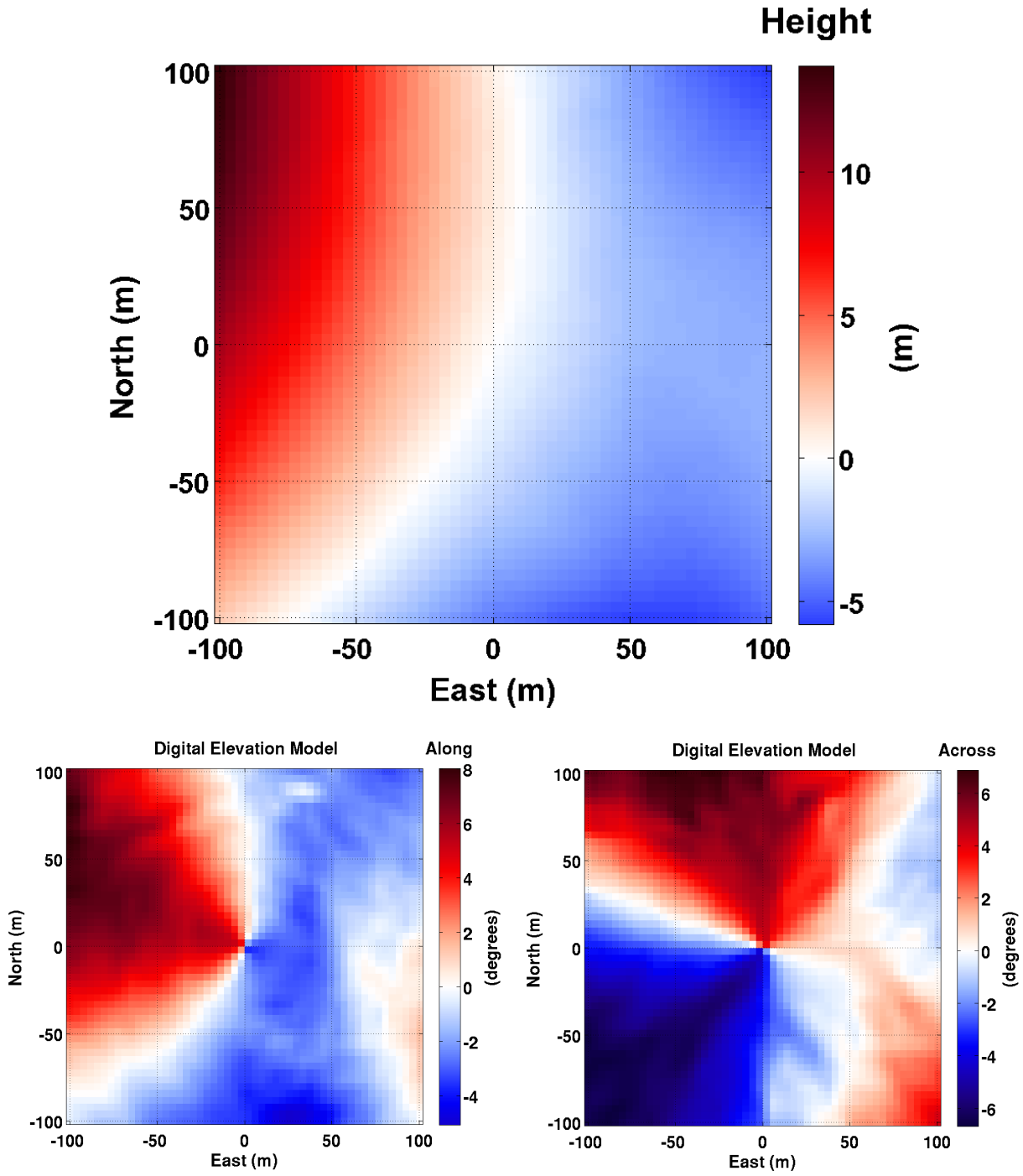

Figure 56: Ground at the alpine site (NWOT); surface height (top), along-track or radial tilting (bottom left), and across-track or azimuthal tilting (bottom right). The GPS is at the origin.

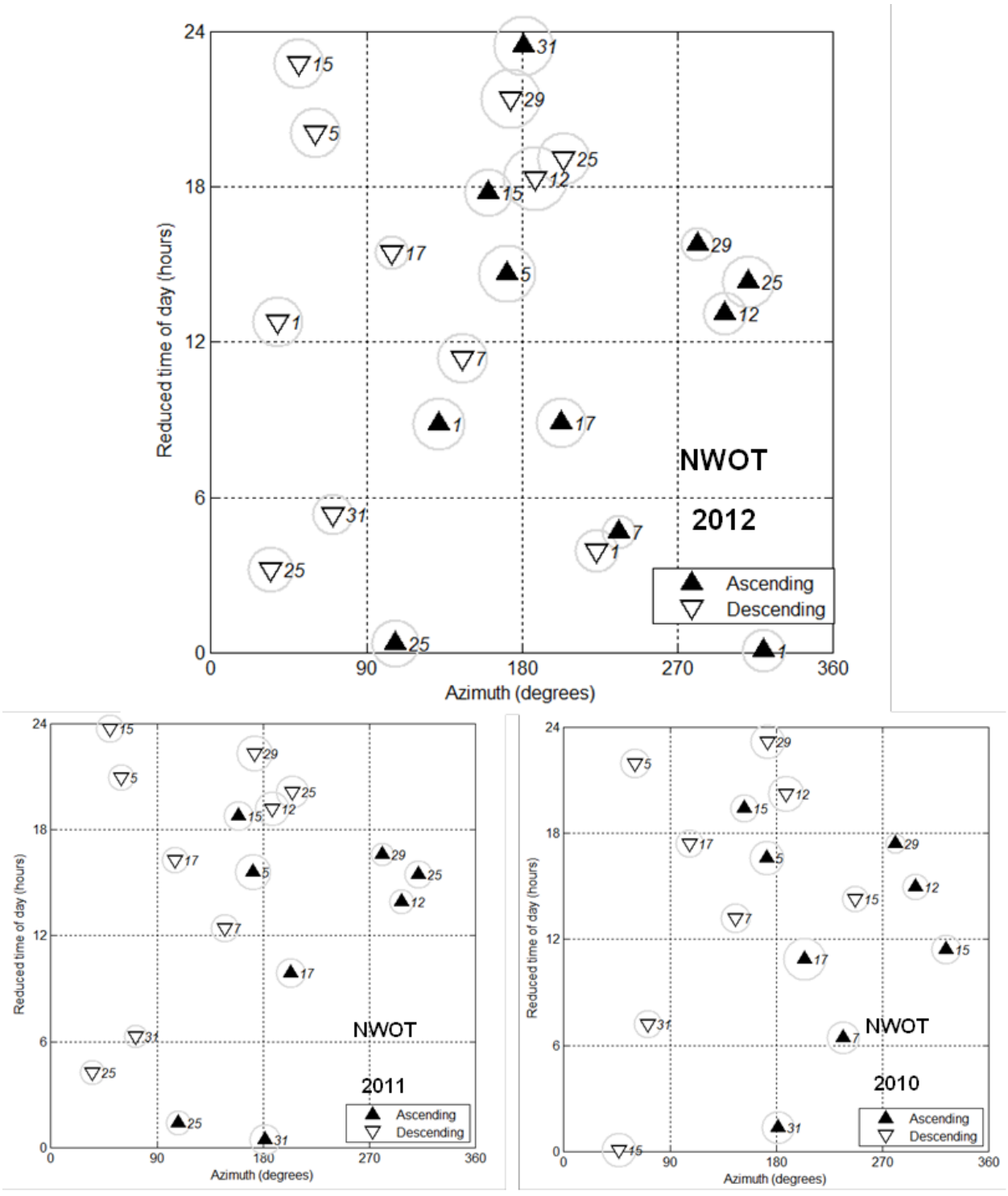

Figure 57: Diagram of satellite track clusters available at least 50% of the year at the alpine site (NWOT), for three water-years: 2012 (top), 2011 (bottom left), and 2010 (bottom right). Italic labels denote satellite PRN number.

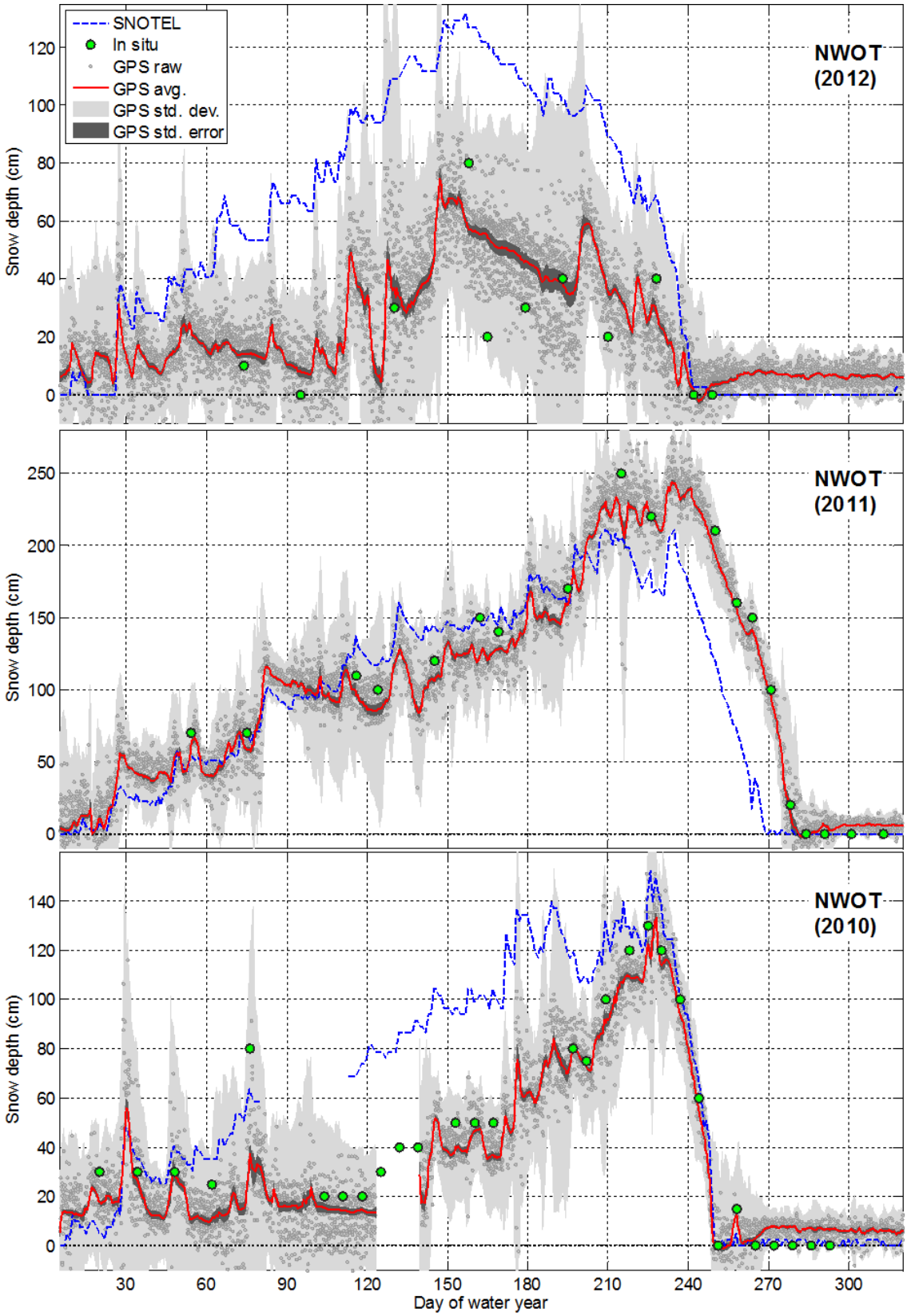

Figure 58: Snow depth measurement at the alpine site (NWOT) for three water-years.

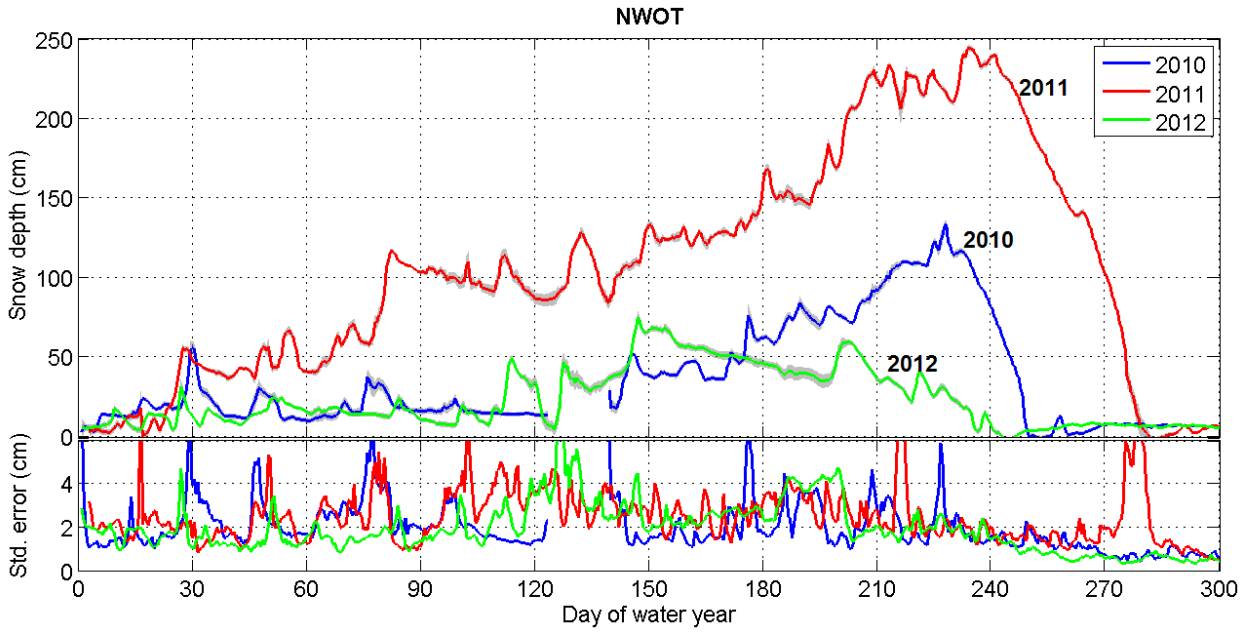

Figure 59: GPS estimates over multiple years at the alpine site (NWOT); the standard error of the mean is shown as gray bands in the top panel and also separately in the bottom panel.

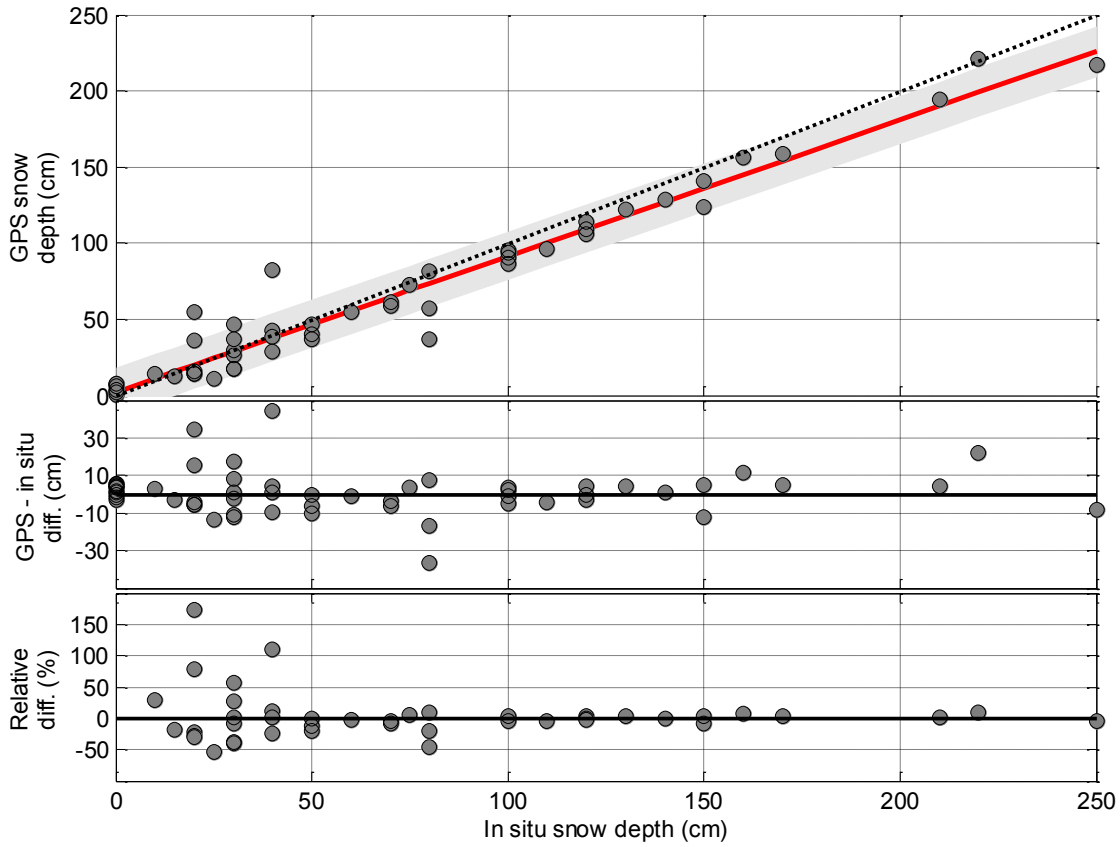

Figure 60: Scatterplot of GPS vs. *in situ* snow depth for all three years at the alpine site (NWOT). A simple linear regression is shown in red, with its 95% observation prediction interval shown as a light-gray band. Post-fit residuals are shown in the middle panel; the bottom panel shows residuals normalized by the *in situ* snow depth value.

Chapter 6: Structural error

Significant though unaccounted for measurement effects will results in structural errors. Systematic effects such as topographical tilting and undulations, medium layering, and volumetric scattering are some examples. Structural error is not expected to be bounded by formal uncertainty, thus leading to possibly unrealistic results. It is also especially dangerous in causing non-zero average errors. Here we single out a few such errors, to the extent that they manifest themselves as reflector heights, thus potentially biasing snow depth retrievals.

6.1 Non-geometrical reflector height defined

Recall that the interferometric phase ϕ_{i} has a geometrical component that reads $\phi_I = k\tau_{\mathrm{i}}$ for a horizontal infinite interface, in terms of the free-space wavenumber k and the interferometric delay $\tau_{\mathrm{i}} = 2H_I \sin e$, where H_I is the reflector height (the subscript I is introduced for consistency). We can rewrite this geometrical interferometric phase as $\phi_I = k_z H_I$ upon definition of the vertical wavenumber k_z:

$$k_z = \partial\phi_I/\partial z = 2k \sin e \tag{87}$$

which can be interpreted as the (elevation angle dependent) vertical sensitivity of the interferometer. Conversely, we find that $\partial\phi_I/\partial k_z = H_I$. This is trivial for the interface phase ϕ_I, but when applied to the other non-geometrical components of the full interferometric phase $\phi_{\mathrm{i}} = \phi_I + \phi_X + \cdots$, the result can be interpreted as an equivalent horizontal reflector height. We now apply this analysis to these other phase components, including some that have been neglected so far.

6.2 HOMOGENEOUS MATERIAL COMPOSITION

The compositional phase ϕ_X has inverse vertical sensitivity $\partial\phi_X/\partial k_z = H_X$ that gives rise to a reflector height H_X born out of the compound reflection coefficient X, which in turn depends on the material composition as well as the antenna response. Thus the total reflector height found via spectral analysis of detrended SNR includes more than just geometrical reflector heights: $H_I + H_X + \cdots$. When we difference out sets of best-fitting sinusoid parameters, between measured and simulated SNR, the compositional reflector height cancels out, provided that the material assumed in the forward model is correct. Changing the material in the simulations will change the estimated reflector height. This issue remains moot for time-differenced reflector heights as long as the material remains the same. Yet when there are changes in medium type – e.g., from snow to soil –, or in medium properties (snow density, soil moisture), compositional reflector height changes may be misinterpreted as geometrical reflector height changes. The differences in compositional reflector height between snow and soil reach up to 5 cm; see Figure 61.

6.3 TOPOGRAPHY

The shape of the reflecting surface is idealized as flat and smooth. Spatially-variable systematic deviations of the mean surface with respect to a plane – called undulations or trend –, may cause the interferometric delay τ_i to fail to conform to its model based on the sine of elevation angle. Considering that any function can be adequately represented in a piece-wise linear fashion by employing sufficiently small pieces, the intervals over which τ_i will remain nearly linear in $\sin e$ depends on the spatial scale of undulations. Severe non-linear variations in τ_i over excessively wide $\sin e$ intervals will be impossible to fit employing the model adopted here. Fortunately this applies on a track-by-track basis, i.e., tracks at azimuths where the surface is nearly planar can

still be fit successfully despite failure of other tracks at azimuths where the surface is severely non-planar. Topographical error depends on the surface tilting along the satellite track. Figure 62 shows that it remains confined to 10% of the snow depth for a 2-m height of the antenna above the ground.

6.4 TEMPORAL CHANGES IN SURFACE SCATTERING HEIGHT

Here we reconsider the interferometric phase ϕ but now permitting reflector height H to be variable; via the chain rule:

$$\partial\phi/\partial k_z = H + k_z Q \tag{88}$$

where the factor $Q = \partial H/\partial k_z = \partial^2\phi/\partial k_z^2$ is a height ramp (or quadratic phase variation). Considering only its time-dependence, it can be obtained as:

$$Q_{\dot{H}} = \dot{H}/\dot{k}_z \tag{89}$$

in terms of the height-rate $\dot{H} = \partial H/\partial t$ and the vertical wavenumber rate $\dot{k}_z = \partial k_z/\partial t$ (this is valid for non-culminating and non-geostationary satellites, otherwise $1/\dot{k}_z = 1/0$). The latter is simply

$$\dot{k}_z = 2k\dot{e}\cos e \tag{90}$$

involving the free-space wavenumber k and the satellite elevation angle e and its rate $\dot{e} = \partial e/\partial t$ (in radians per second, not degrees per second). Putting everything together, we find a temporal reflector height component $H_{\dot{H}}$ of the form[2]

$$H_{\dot{H}} = k_z Q_{\dot{H}} = \dot{H}\tan e\,/\dot{e}. \tag{91}$$

[2] If SNR were to be simulated under non-zero height-rate conditions, it would involve the integrated $\Delta\tau_\mathrm{i}$ rather than the instantaneous τ_i interferometric delay (where zero-subscripted quantities are piece-wise constant):

$$\Delta\tau_\mathrm{i}(t) = \tau_\mathrm{i}(t_0) + \int_{t_0}^{t} \dot{\tau}_\mathrm{i}(t)\,\mathrm{d}t = \tau_\mathrm{i}(t_0) + 2H_0\Delta\sin e + 2\dot{H}_0\Delta t\sin e\,.$$

The dependence on $\dot{e}$ implies that simultaneously rising and setting satellites will respectively under- and over-estimate geometrical reflector height, i.e., will experience temporal reflector heights $H_{\dot{H}}$ with opposite sign. Larson et al. (2013) found meter-level $H_{\dot{H}}$ for a GPS tide gauge site exhibiting 7-m daily amplitude tides and maximum height-rates of $\dot{H} \approx 1.5$ m/h. The relevant temporal scale corresponds to the duration of an individual track, i.e., the non-zero height-rate has to sustained for 1-2 hours to have an impact.

For snow, this effect is not nearly as large – normally about 5 cm –, but we did detect its presence during the heaviest snow precipitations. Indeed, we have rejected reflector height outliers of opposite sign for ascending and descending tracks, which is strong evidence for the present source of the error. During those events, total reflector height is steadily decreasing over the entire site – from the perspective of the GPS, it is as if the reflecting surface were uplifting as a whole. A rule-of-thumb to predict the magnitude this reflector height error (in meters) is ¾ of the height-rate (in meters per hour), i.e., $H_{\dot{H}} \approx \dot{H} \times 45$ min. So a 10-cm/h rate yields a ~ 7.5-cm error (±10%); see Figure 63. As long as the site has visibility to both ascending and descending tracks, the site daily average should be unaffected in accuracy, affected only in precision, by this error.

6.5 ANGULAR CHANGES IN VOLUMETRIC SCATTERING DEPTH

In the snow-covered period, reflector height retrievals are assumed to correspond to the upper surface of the snow, i.e., the snowpack volumetric scattering center is assumed to coincide with the air/snow interface. This assumption seems to be well validated by external *in situ* measurements. Yet it is well known from the radar literature (Ulaby et al. 1986) that dry snow is very transparent at L-band; comparisons between InSAR (microwave) and LIDAR (optical) show up to 100-m bias (Rignot et al. 2001). So there is an apparent contradiction between GPS

and conventional radar studies. The fact that such a non-zero geometric/electromagnetic reflector height bias has so far gone undetected in experiments must mean that it is smaller than the dispersion of both GPS and *in situ* measurements.

Volumetric scattering arises because snow, although homogeneous at the macroscopic scale, exhibits random inhomogeneities at the microscopic level – it is a mixture of inclusions (ice particles) embedded in a host medium (air). This phenomenon happens even in the absence of internal layers – either discrete interfaces or continuously-varying stratification. It can be accounted as follows.

Recall that the forward model involved the complex-valued coherence $\gamma = \gamma_S \cdot \gamma_V \cdot ...$, which can be expressed as the product of various processes: surface scattering γ_S, volumetric scattering γ_V, etc. The degree of coherence $|\gamma| = |\gamma_S| \cdot |\gamma_V| \cdot ...$ never strengthens, only degrades, as each factor is smaller than unit. The interferometric phase $\phi_\mathrm{i} = \phi_\gamma = \phi_S + \phi_V + \cdots$ accumulates all the phase contributions. Up till now, we have accounted only for surface scattering $\phi_S = \phi_I + \phi_X ...$ (interface plus composition contributions) in the forward model; volumetric scattering can be readily appended.

Here we consider the simplest model of coherent volumetric scattering, namely, specular incidence and scattering off a half-space within which power decays exponentially with depth (Cloude 2009, eq.(7.14)):

$$\gamma_V = \left(1 + \imath k_z^2/(4k\sigma_\mathrm{e})\right)^{-1}. \tag{92}$$

It is parameterized in terms of the mean extinction σ_e (in units of m^{-1}). Although developed for polarimetric interferometric synthetic aperture radar (PolInSAR) applications, where backscattering and mono-static configuration prevails, this model seems applicable to bi-static configurations as found in GPS-MR.

Calculating the inverse vertical sensitivity of the volumetric phase $\phi_V = \arg(\gamma_V)$ we find a volumetric reflector height of the form:

$$H_V = \frac{\partial \phi_V}{\partial k_z} = -\frac{\sigma_\mathrm{e} \sin e}{k^2 \sin^2 e + \sigma_\mathrm{e}^2}. \tag{93}$$

Figure 64 illustrates the formula. Although this result may or may not be quantitatively correct, it is qualitatively reasonable, in that (*i*) it predicts a negative volumetric height, i.e., the volumetric scattering center is buried under the surface; and (*ii*) it demonstrates how volumetric scattering degenerates into surface scattering as the satellite converges to grazing incidence.[3]

6.6 CONCLUSIONS

In previous chapters we saw how inversion errors depend on a multitude of interacting effects. Estimated uncertainty will bound the actual error provided that the postulated functional and stochastic models match the measurement-generating processes. This applies to the average inversion performance, as long as individual realizations, though random, are drawn from the postulated distribution. A failure in the error/uncertainty bounding indicates a mismatch between models and reality, not an irreducible flaw in the underlying inversion machinery.

[3] For completeness, the volumetric height ramp reads:

$$Q_V = \frac{\partial H_V}{\partial k_z} = -\frac{\sigma_\mathrm{e}^3 - 3k^2 \sigma_\mathrm{e} \sin^4 e}{2k(k^2 \sin^4 e + \sigma_\mathrm{e}^2)^2}.$$

At grazing incidence we find $Q_V|_{e=0} = -1/(2k\sigma_\mathrm{e})$, whereas for the temporal height ramp we have instead $Q_{\dot{H}}|_{e=0} = \dot{H}/(2k\dot{e})$, so the two effects seem amenable to separation.

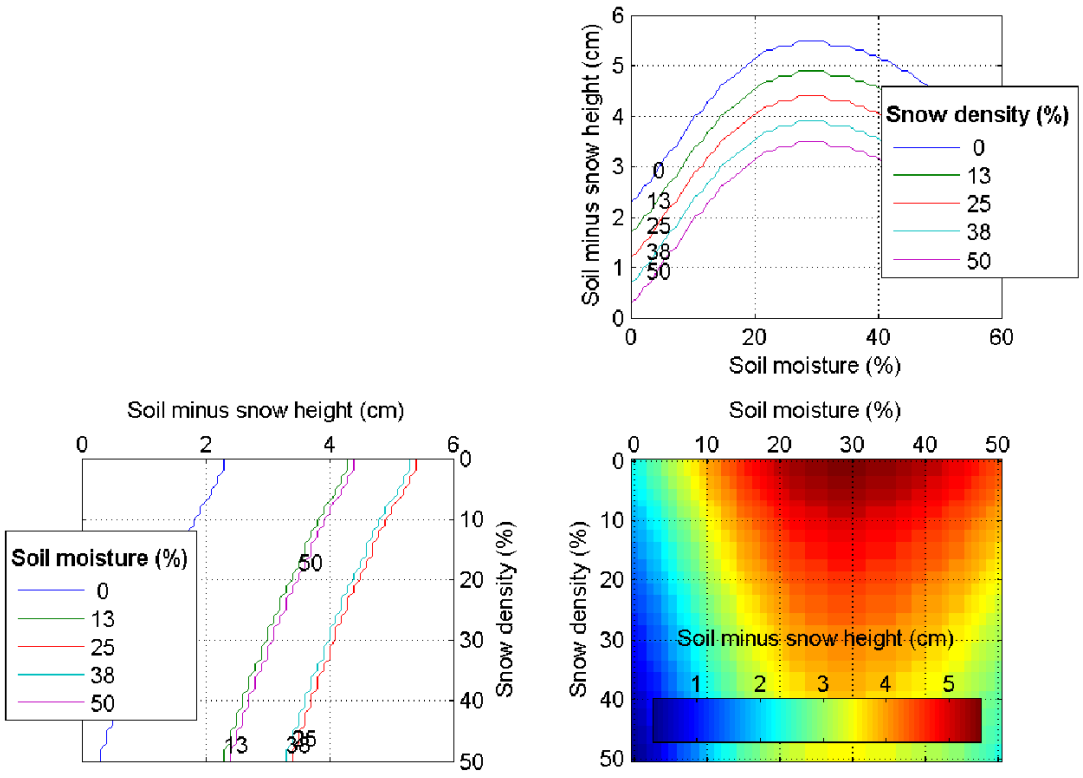

Figure 61: Compositional reflector height variation with snow density and soil moisture for a 2 m height of the antenna above the surface.

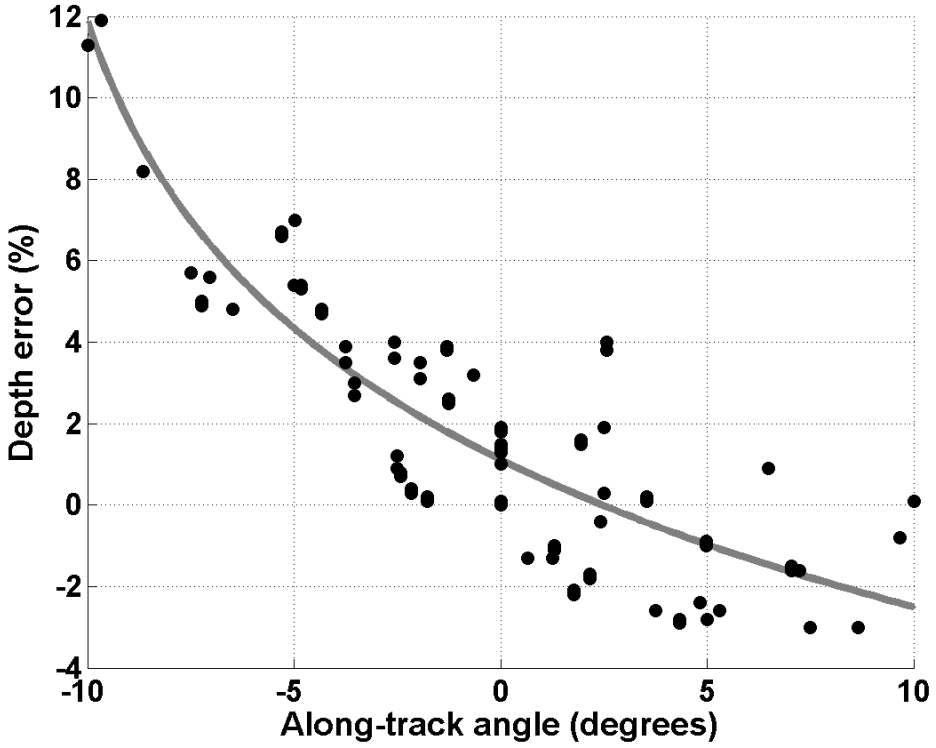

Figure 62: Snow depth error for varying topographical tilting.

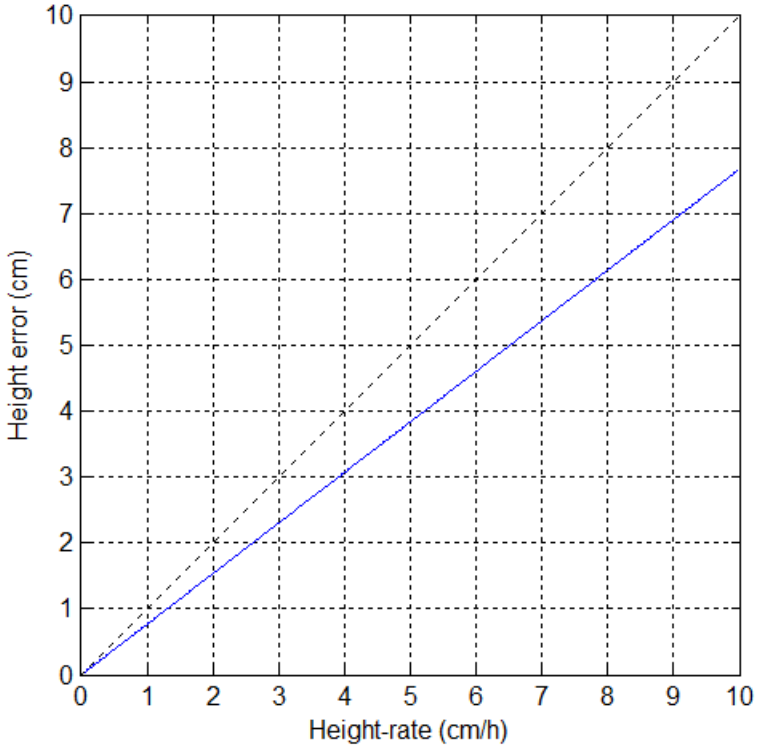

Figure 63: Relationship between temporal reflector height error and geometrical reflector height rate.

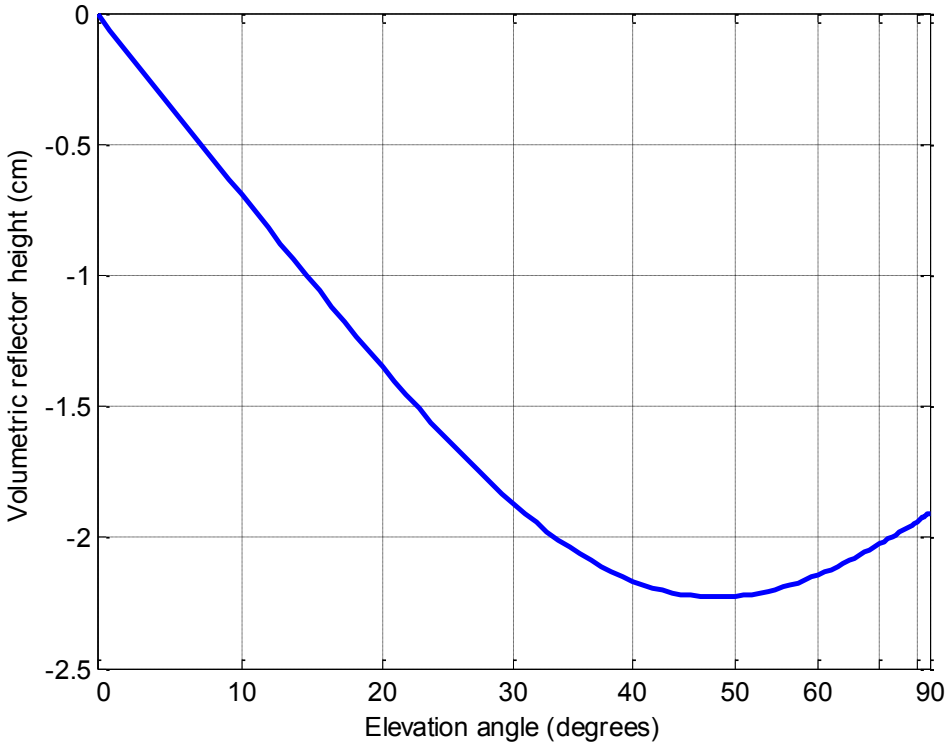

Figure 64: Relationship between volumetric height and satellite elevation angle.

Chapter 7: Conclusions

In this work I offered a contribution for ground-based monitoring of snow depth using the technique of GPS multipath reflectometry (GPS-MR). GPS-MR exploits the simultaneous reception of reflected signals in conjunction with the direct signal transmitted from a GPS satellite. It relies on the coherence of such signals to produce constructive and destructive interference fringes. This dissertation provides three chapters that detail a full forward model, development of an inversion method, and application of that method to data from three field sites for 1-3 years. A final chapter examines several key factors related to the accuracy of the method. Here I summarize the main findings of my dissertation.

7.1 Forward modeling

Simulations based on the forward model enabled the study of a number of aspects related to GPS multipath in SNR observations, pertinent to either or both positioning and reflectometry near-surface applications. The intent behind bridging the two fields was to foster their cross-fertilization, such that developments in multipath mitigation for positioning may be leveraged for multipath exploitation in reflectometry, and vice versa.

Here are some individual remarks concerning the general forward modeling efforts. The more fundamental direct and reflected signals can be manipulated to define more useful quantities of interest in reflectometry and positioning applications, respectively the interferometric and error signals (their power, phase, and delay). Coherence was shown to play a crucial role, in that it is a pre-requisite for multipath effects to be observed in SNR and carrier-phase measurements. Incoherent power may additionally affect pseudorange measurements, depending on the code discriminator employed. The incident signal polarization ellipticity is

negligible in GPS-MR snow depth retrievals. Code discriminator can be neglected for the small interferometric delays typically found in near-surface installations and near-grazing incidence.

I found that although metallic surfaces may produce a strong reflection electric field, they result in a weak reflection voltage as a consequence of the antenna/surface polarization mismatch (LHCP vs. RHCP). Because of the lack in diversity with respect to the direct signal – small interferometric delay and Doppler, same sense of polarization, and similar direction of arrival –, geodetic GPS antennas offer little impediment for the intake of near-grazing reflections off dielectric natural surfaces. Reflection phase includes a compositional component in addition to the geometrical or propagation-related component. SNR will deviate from an idealized sinusoid form because of the incident power trend, random surface roughness, antenna gain, and surface material composition. SNR physical modeling agreed better for longer public codes (like L2C). The degradation of SNR for longer encrypted codes (P(Y)) is attributed to codeless tracking losses, whose trend can be accounted for empirically. The degradation of SNR for shorter public codes (C/A) is attributed to self-interference, which would seem to be more difficult to model. Self-interference repeats with the satellite constellation (~ every sidereal day) so it can be misinterpreted as multipath.

For multipath reflectometry purposes, it is important to highlight that coherence is a fundamental physical limit, which might hamper the sensitivity of observations with respect to any other medium parameter beyond a certain phase stability threshold. Surface random roughness is only one source of decoherence. Conclusions were reached for practical purposes, too: tipping the antenna allows neglecting antenna gain pattern in dielectric surfaces; it also helps with self-interference in shorter modulation codes. Longer public codes (L2C) signal allow for using upright installations, with the added benefit of an increased number of visible satellites. On

the other extreme, upside-down antennas do not guarantee reflection-only reception, especially near grazing incidence. For soil moisture sensing with geodetic GPS antennas, the polarimetric antenna phase difference was confirmed decisive in translating changes in reflection polarization ellipticity into phase-shift retrievals. For snow depth sensing, topographical reflector height bias is constant for GPS (except in satellite maneuvers).

On the other hand, for positioning applications, it was found that the pervasive assumption of LHCP-only reflection leads to severe underestimation of the carrier-phase error near grazing incidence off dielectric surfaces. The coincidence (or lack thereof) of peaks and troughs across code and carrier multipath errors is dependent on the surface material and on the polarimetric antenna phase difference. For deformation monitoring, it seems preferable to locate the antenna lower – perhaps flush with the soil – to avoid time-dependent apparent displacements. In contrast, for reference frame realization it would be better to locate the antenna higher ($> \sim 2$ m), to help average-out carrier-phase multipath error. The presence of a coherence-reducing environment (e.g., randomly-shaped objects, tall grass, etc.) in the antenna surroundings would appear to be a valid carrier-phase multipath mitigation strategy. Metallic horizontal surface is less harmful than bare ground at near-grazing incidence with geodetic-quality antennas (not necessarily so at near-normal incidence).

7.2 INVERSE MODELING – FORMULATION

In developing the inverse model, the unknown conditions in the antenna surroundings, such as the amount of snow, were parameterized in terms of a few biases to the *a priori* values assumed for the physical parameters. These biases can then be embedded in the forward model, modifying the interferometric power and phase and the direct or noise power. The parameters set had to include not only the one immediately of interest (namely, reflector height for snow depth

sensing) but also a few nuisance parameters (such as phase-shift) without which observations cannot be adequately fit. SNR observations were partitioned in terms of ascending and descending satellite tracks, within which the unknown biases parameters can be assumed constant. For an individual track, one cannot separate physical effects that impact observations in a similar manner. These nearly linearly dependent parameters can sometimes be separated combining multiple track instances in post-inversion processing.

As the forward model relates observations and parameters in a non-linear manner, there is a need to obtain reasonable starting values for the unknowns though a preliminary global optimization, without which the final local optimization is not guaranteed to converge to the optimal solution. This can be achieved through a polynomial/spectral decomposition of SNR observations into a trend plus fringes. Applying the same decomposition to simulated and measured observations then differencing the results cancels out the effects already accounted for in the forward model, such as antenna gain patterns and the height of the antenna above the ground. The sensitivity of observations to each unknown parameter constitutes the inversion functional model; a good parameterization yields distinct sensitivities for each parameter. The stochastic model scales these sensitivities non-uniformly employing the noise characteristics as experienced by observations.

I simulated the inversion performance when faced with random noise. It quantified the expected errors (defined as the difference between retrievals and true or known values for the simulated parameters) as well the uncertainty (and how well such error-bars bracket the error itself). The parameter of interest (reflector height bias) was found to be difficult to determine when the total reflector height (*a priori* value minus estimated bias) was smaller than ~ 0.5 m. This followed from the ambiguity in the polynomial/spectral decomposition, as the low-

frequency fringes resembled the low-order trend in SNR observations. Secondarily, the uncertainty was also found to be impacted by the amplitude of fringes, as the peaks and troughs became less distinct against the random noise background. Increasing the number of nuisance parameters (e.g., maximum polynomial degree) further degraded the performance; this indicated the importance of *a priori* antenna gain patterns to aid with the SNR detrending.

I also investigated the directional dependence of results. Observations at lower elevation angles were confirmed to be more important, as immediately apparent upon inspection of the greater amplitude of interference fringes there. Higher elevation observations were found to become increasingly more important for decreasing total reflector heights and for increasing number of unknown parameters, as these observations help mitigate the above-mentioned ambiguity between SNR trend and fringes. I was also able to establish the sensing footprint upon propagation of the *a posteriori* parameters covariance matrix. The resulting uncertainty in the full phase bias function – containing constant and linear terms (phase-shift and reflector height, respectively) – exhibited a dip for lower elevation angles, indicating the satellite directions where the interferometric phase is better determined; the best-determined direction was named the peak elevation angle ($\sim 10°$). A sensing weight (between 0 and 1) was defined as the normalized reciprocal of uncertainty, which when plotted versus the center of the first Fresnel zone at each elevation angle, indicated peak radial distances near ~ 15 m, with a longer far tail and a shorter near tail (respectively regions beyond and closer than the peak distance). The implications for *in situ* data collection are evident: sample more intensely around peak distance, less so in the immediate vicinity of the GPS, tapering it off gradually away from the GPS. As a caveat, these conclusions are not necessarily valid for antenna setups other than the one considered here, e.g., tipped instead of upright antenna, or installations much above the surface, $H \gg 2$ m.

7.3 INVERSE MODELING – APPLICATION

In applying the inverse model to snow sensing, I illustrated how SNR measurements can be disrupted in many different ways, including but not limited to secondary reflections, precipitation, and instrument-related issues. Even if these error sources have a definite cause, as long as they vary erratically in the span of a GPS track, they cannot disrupt the inversion procedure. On the other hand, error sources manifesting exactly as a reflector height (i.e., with similar observation/parameter sensitivity), such as the largest scale ground topography, are not distinguishable from snow depth on a single track basis; these may only be separated given redundant track results. In between these two extreme regimes, it was found error sources that are partially random and partially systematic during a GPS track. This was the case for some types of instrument issues and medium-scale surface deviations as well. They failed to conform to a linear model in terms of $\sin e$ but did exhibit non-negligible auto-correlation over elevation angle. A fourth source affecting the quality of SNR observations leads to the extinction of the interference fringes. Although it affects the quality of snow depth retrievals, this is not an additive error source such as the ones above, rather it is a consequence of genuine physical processes impacting the interferometric power.

I explored the satellite coverage in terms of rising/setting azimuth and time of the day as well. The thousands of individual track results were organized in only ~ 20 clusters based on the repeatability of either of these two variables. Displaying the cluster locations in an azimuth/time-of-day diagram provided an inventory of the coverage available at a site, which was a consequence primarily of the sky visibility, as dictated by, e.g., obstruction by trees.

I proceeded to show how these track clusters can serve as the basis for quality control (QC) of results, as each cluster forms a more statistically homogeneous unit. Four QC tests were

demonstrated: statistical degree of freedom, RMS of residuals, peak elevation angle, and reflector height uncertainty. Throughout, the guiding principle was comparing day-to-day independent track estimates to their long-term tendency; outliers were detected vis-à-vis the typical dispersion. Statistically robust estimators (e.g., median instead of mean) were found particularly effective in obtaining these first- and second-order statistical moments. None of the four QC tests had sufficient power alone, but in conjunction they were found adequate.

The general analysis concluded with the formation of inter-cluster site-wide daily snow depth averages. It required harmonizing results across different clusters by removing topographical effects, which impact primarily reflector heights (up- vs. down-hill) and secondarily also snow thickness values (ground hills and depressions, or concave vs. convex terrain). The fact that not all tracks are of the same quality (because of obstructions, etc.) was accounted for employing variance scale factors for the inversion formal uncertainty. I also emphasized the distinction between standard error of the mean and standard deviation of observations, to be kept in mind when comparing independent results.

After these generic developments, I presented site-specific results, validating GPS snow depth retrievals against independent *in situ* measurements collected at grasslands, forested, and alpine locations. At the forested site (RN86) snow depth reached ~ 1.5 m. *In situ* samples were collected in 9 visits spanning half a year in a spatially dense manner (20-to-150 values per sample). The assessment at RN86 yields a correlation of 0.984 and an RMS error of 6.8 cm; the GPS is found to under-estimate *in situ* values by ~ 18 ± 13%. The alpine site (NWOT) received more snow (~ 2.5 m) and the *in situ* data collection had greater temporal coverage (90 visits spanning three years) albeit less intense spatially (one value per sample). The snow depth assessment statistics at RN86 were similar to those observed at NWOT: very strong correlation

(0.980) and sub-decimeter RMS (7.8 cm); the GPS under-estimation was smaller and also better determined, ~ 11 ± 3%, although the footprints are not overlapping. The GPS daily site averages were found effective is mitigating random noise without unduly smoothing the sharp transitions as captured in, e.g., new snow events.

The weakest link in now depth monitoring with GPS-MR is the determination of the site average zero-depth line. While cluster-wise topographical heights can be obtained reliably as the median over the a long snow-free period, the site average bare-ground height relies on much fewer observations, normally the 5^{th} percentile of the daily averages over the same period. The fact that reflectors height retrievals continue to vary by 5-10 cm – one order of magnitude larger than the ~ 1-cm precision of the site daily averages – even long after snow has completely melted, indicates the presence of unaccounted for systematic effects that may be contaminating snow depth retrievals by the same magnitude throughout the season. This error still amounts to only few percent of the total snow depth sensed.

7.4 STRUCTURAL ERRORS

In the last chapter I considered some physical effects that are currently neglected in the forward/inverse model. Although not bounded by the formal uncertainty, they ought to be considered in a realistic error budget. The vertical wavenumber k_z was introduced to quantify the vertical sensitivity of the GPS as an interferometer. It vindicated our inversion formulation using $\sin e$ as the basis for a polynomial expansion of the unknown biases, revealing that constant, linear, and quadratic terms are coefficients in powers of k_z. In converse, non-geometrical reflector heights were defined as the sensitivity of interferometric phase components to the vertical wavenumber. Compositional, topographical, temporal, and volumetric reflector heights were examined, all of which might be present in best-fitting reflector heights. I found error

magnitudes in the range of tens of centimeters, which should be considered as an approximate figure of merit, not a definite comprehensive error budget. The evaluation of the volumetric scattering contribution was facilitated by our forward formulation in terms of complex coherence, thus bringing us full circle to the forward model introduced early in this work.

8. Future work

Here I make some recommendations for future work that could further advance the progress made in this research.

8.1 FORWARD MODELING

The theoretical forward model and its computational implementation in MATLAB could be extended in a number of ways. I sought to provide gain patterns for a number of antenna models used in GPS networks, which allowed comparing their suitability for positioning and reflectometry applications; gains for additional antennas could be easily incorporated. It is also possible to add additional carrier frequencies (e.g., L5) and modulations (e.g., BOC), from GPS and other GNSS as well. The effect of random surface roughness on phase needs to be implemented, as I have implemented only its power effect. The simulator could also incorporate different formulations for the SNR estimator (Falletti et al. 2011); I have neglected the under- or over-estimation of SNR by non-ideal GPS receivers.

Here I have considered only stationary surfaces and receivers. Dynamic scenarios, as in tidal waters, would be not just a time-succession of static scenarios. It would require accounting for the interferometric Doppler, not only in the code modulation (sinc factor) but more importantly in terms of the accumulated interferometric delay. Although I have considered temporal and volumetric reflectors heights, the simulator presently does not incorporate these structural errors in producing SNR observations.

Extended surface geometry – beyond the current horizontal model –, accounting for tilted (though still planar) surfaces as well as large-scale undulations (with potentially multiple simultaneous reflections) has been developed and will be reported at a later time.

8.2 INVERSE MODELING – FORMULATION

In hindsight one parameter that would have been worthwhile to include in the inversion of SNR observations is a constant height ramp or quadratic phase bias coefficient, of the form $\phi_B^{(2)} \sin^2 e = Q_B k_z^2$ where $\phi_B^{(2)} = 4k^2 Q_B$. It could have prevented reflector height outliers of opposite sign during the heaviest snowfalls. Also, instead of using powers of $\sin e$ as polynomial basis for the expansion of the phase and power biases, the vertical wavenumber $k_z = 2k \sin e$ would have been a more defensible choice.

The quadratic phase parameter in itself would open up the possibility of new retrievals. At shorter-periods (1-2 h) there is a dependence on precipitation rates, through $\dot{H}$ in $Q_{\dot{H}} \approx \dot{H}/(2k\dot{e})$. Beyond day-long periods, the mean extinction σ_e in $Q_V \approx -1/(2k\sigma_\mathrm{e})$ is a proxy for soil moisture and snow density. It remains to be seen how well Q could be decorrelated from the other phase and power biases, something dictated by how distinct its observation/parameter sensitivity is from that of the other parameters. The topographical effect in Q could presumably be separated as done for reflector heights H, i.e., assuming it constant for a single day.

If attempting to exploit the power biases for reflectometry, it would have been better to split the direct or noise power bias K in two:

$$\mathrm{SNR} = \left(1 + P_\mathrm{i} + 2\sqrt{P_\mathrm{i}} \cos \phi_\mathrm{i}\right) P_\mathrm{d}^\mathrm{C} K^\mathrm{C} / P_\mathrm{n} + P_\mathrm{c}^\mathrm{I} K^\mathrm{I} / P_\mathrm{n}. \tag{94}$$

As the SNR trend involves both components of the direct power, $P_\mathrm{d} = P_\mathrm{d}^\mathrm{C} + P_\mathrm{d}^\mathrm{I}$, the incoherent composite power $P_\mathrm{c}^\mathrm{I} = P_\mathrm{d}^\mathrm{I} + P_\mathrm{r}^\mathrm{I}$ could contaminate coherent power bias retrievals if left unaccounted for. It needs to be checked whether or not the inversion would remain well-conditioned after the inclusion of K^I, as it could be nearly linear dependent with K^C and other parameters, especially for shorter tracks.

8.3 INVERSE MODELING – APPLICATION

Inter-cluster combinations could be extended to resolve azimuthal asymmetries by estimating one or a few azimuthal gradients instead of just the site average as done in this work. Given the distribution of track clusters, the site average would improve in accuracy as it becomes less biased towards the azimuthal regions more densely sampled, although its precision would degrade because of the greater number of parameters. Application to snow depth is useful as snow is known to be spatially variable as a consequence of a variety of controls on its deposition and ablation. Applied to reflector heights, it could offer a map of the ground around the GPS, which could be accurate enough to aid in correcting for the topographical structural error in snow depth retrievals. The two applications could be assessed and validated with azimuthal *in situ* snow and topographical datasets already collected.

Inter-cluster combinations could also be extended to parameters other than reflector heights or snow depths, such as phase shifts and power biases. They would have to be harmonized by removing the topographical effect, although it is unclear whether that would be an additive error as for reflector heights or a more complicated proportional effect.

Inter-cluster combinations could be generalized to deal with multiple parameters, e.g., all interferometric phase biases (phase-shift, reflector height) or all interferometric power biases (zeroth-, first-, and second-order) in conjunction. Although a multivariate combination is admittedly more complicated, it is advantageous in that it accounts for the significant correlation existent between similar parameters – e.g., phase-shift would improve the precision of reflector heights and vice-versa.

Alternatively, one could retain the simpler univariate combinations but instead of applying it to the multiple polynomial coefficients underlying the full phase or power bias

functions, it could be applied to the scalar resulting from evaluating these functions at a fixed elevation angle, e.g., the median minimum-variance elevation angle (which varies slightly from day to day). These different types of daily site averages could be integrated to shed light on the so-called model identification problem, i.e., to discern the type of scattering target, between, e.g., snow and ground and vegetation.

Applications to other worldwide locations, such as Greenland and Antarctica, could be easily accomplished, assuming minimal impact from the latitude dependency in the satellite visibility. Furthermore, although I have explored only reflector heights for snow depth sensing, many other parameters are produced as a by-product of the SNR observation fitting. Therefore the present forward/inverse approach should remain applicable for the study of targets other than snow. For example, phase-shift is known to be sensitive to soil moisture (Small et al. 2009); the power biases (noise and interferometric) are known to be sensitive to temperature, vegetation, snow density, etc. Additionally, the RMS of residuals could potentially be used as a sensing variable. This proposition is based on the hypothesis that SNR errors are partially instrumental and partially physical. The former is always present while the latter would change with the scattering conditions, particularly with the reflection coherence.

Time series of SNR residuals – not just their summarizing statistics – could also be useful in themselves. For example, SNR observations are known to be affected by space weather. Ionospheric scintillations – rapid fluctuations in direct power due to irregularities in the plasma enveloping Earth's magnetosphere – are notorious for causing GPS receivers to lose lock, thus leading to missed measurements. In a complementary fashion, ionospheric monitoring could benefit from reflectometry developments, in separating these two sources in SNR (Bishop et al., 1985). Candidate ionospheric events could be detected based on the QC tests, then corroborated

or rejected based on whether or not they are observed at widely spaced simultaneously operating receivers.

8.4 STRUCTURAL ERRORS

Further investigation is warranted towards more elaborate models of coherent volumetric scattering (Cloude, 2009), such as random-volume-over-ground (RVOG) for snowpacks and oriented-volume-over-ground (OVOG) for vegetation, the latter of which is able to distinguish the polarimetric response of, e.g., vertical corn stalls versus horizontal leaves.

Tropospheric refraction has been tacitly assumed negligible. It would encompass both the interferometric propagation delay and angular changes in the signal direction of arrival. These effects become more important for larger reflector heights and smaller satellite elevation angles. It needs to be assessed what are the cutoff height and elevation angle values beyond which refraction becomes significant in GPS-MR.

I also neglected internal layers, both discrete interfaces (such as snow/ground) and continuously-varying stratification (e.g., from lower to higher snow density). This might have been one cause of the difference in performance between large and small snow depth retrievals, as thin snow layers would be more transparent to the underlying ground. I would claim that accounting for layering while neglecting internal roughness would be even more erroneous than neglecting it entirely, as it may severely over-estimate the coherent contribution of secondary reflections. Media layering with rough interfaces (not just rough top) should be addressed in conjunction with random volumetric inhomogeneities, all of which conspire to annihilate coherence.

8.5 OTHER EFFORTS

Future work should include the assessment of non-L2C SNR observables for snow sensing. Ideally this would be tackled employing SNR recorded simultaneously at different signals. The ability of inter-cluster combinations to mitigate noise improves the prospects of utilizing these legacy GPS signals (C/A and P(Y)). Although noisier than L2C, they are broadcast by three times more GPS satellites. As the latter gain partially compensates for the former loss, the daily site averages might be acceptable, as more frequent sampling improve noise mitigation. If proved satisfactory, this development would allow extending the GPS time series into the past by leveraging historical data. The decadal records produced would be useful in climatic investigations, such as the receding end of the snow season. On the engineering side, one would have to address the issue of potential inter-code, inter-frequency, and inter-system biases present in GPS-MR (or GNSS-MR) reflector height retrievals, e.g.:

- inter-code, intra-frequency: L2C vs. L2-P(Y);
- intra-code, inter-frequency: L2-P(Y) vs. L1-P(Y);
- inter-code, intra-frequency: L1-P(Y) vs. L1-C/A;
- inter-system: GPS, Glonass, Galileo, Beidou.

Multiple carrier bands also allow for wide-laning attempts in GPS-MR, by which longer, meter-level wavelengths are synthesized by the linear combination of existing carriers (Lowe et al. 2007; Martin-Neira et al. 2002). These derived carriers approach the UHF band and would enable sensing targets that are incoherent at L band.

Applications for improved GPS positioning, akin to Bilich et al. (2008), would seem feasible. The estimated parameters inverted from SNR measurements could be utilized to forward-model and subsequently correct multipath errors in code pseudorange and carrier-phase

measurements. This would be restricted to base stations located in relatively simple multipath environments, given our assumption of locally horizontal surfaces and a few constant unknown biases per satellite track. One could also envision monitoring monument vertical stability based on long time series of reflector heights; this would help ensure the position time series represent tectonic movements and crustal deformations rather than local effects. Even in the realm of mere simulations (no measurements), the forward model here developed is more realistic than most used in assessing the impact of multipath errors in positioning.

In terms of hardware developments, there is a need for more versatile research-grade instruments to test the scattering forward models and assist in the development of inverse models. The black-box geodetic equipment utilized in this work outputs only the peak-power or maximum-correlation SNR, while a multi-correlator waveform-outputting receiver could provide the full correlation-versus-delay profile, sometimes even complex-valued (power and phase). One could adapt existing receiver designs, especially software-defined ones, such as those developed for transmitted signal deformation monitoring, incoherent reflectometry, and ionospheric scintillation monitoring. More elaborate antennas would also be helpful for experimentation, such as dual- and quad-polarized antennas – circular, L/R, and linear, V/H –, electronic beam-steering antenna arrays, vertically-displacing antenna platforms, etc. GPS ground-based transmitters ("pseudo-lites") and retransmitters would allow for more flexible illumination configurations. Experimental efforts in utilizing multiple receiver/antenna pairs to mitigate multipath could be leveraged for reflectometry purposes. At the same time there is a demand for simpler and less expensive equipment (Chen et al. 2013), which would enable a wider popularization of GPS-MR and installations dedicated primarily for reflectometry rather than positioning applications, especially at worldwide locations where theft and damage is a

concern. Even considering only geodetic-quality and surveying-grade GPS equipment, the LHCP antenna phase pattern needs better characterization; it could be measured in an anechoic chamber or simulated with antenna modeling software. Efforts are necessary in standardizing SNR measurements, as RINEX "S" observables are currently far from being receiver-independent, unfortunately (Bilich et al. 2007). This issue could be tackled comparing the output of different receivers connected via a splitter to a common live antenna, hardware simulator, or record/playback system.

One last issue that would require more effort is the localization of the important scattering regions in the formation of an individual specular reflection. Although I have determined elevation angle dependent sensing weights, on the horizontal plane each individual weight has been assigned to the center of the first Fresnel zone. This solution could be improved dispersing each weight gradually over the surface instead of assuming a point-like distribution. The final solution would be a convolution of such a spatial averaging kernel with the proposed elevation dependent weights. The determination of the important scattering regions for observables other than SNR (i.e., code pseudorange) also remains an open issue.

References

Aloi D, Van Graas F (1999) Analysis of the effects of Earth-surface based multipath reflections on GPS code-phase measurements. Proc ION AM. Cambridge, MA, pp 609–618.

Anderson KD (2000) Determination of Water Level and Tides Using Interferometric Observations of GPS Signals. J Atmos Ocean Tech 17:1118–1127. doi: 10.1175/1520-0426(2000)017<1118:DOWLAT>2.0.CO;2

Armstrong RL, Brodzik MJ, Knowles K, Savoie M (2005) Global monthly EASE-Grid snow water equivalent climatology. http://nsidc.org/data/nsidc-0271.html.

Armstrong RL, Brun E (2008) Snow and Climate – Physical Processes, Surface Energy Exchange and Modeling. Cambridge University Press, p 222 pp.

Barnett TP, Adam JC, Lettenmaier DP (2005) Potential impacts of a warming climate on water availability in snow-dominated regions. Nature 438:303–309.

Beckmann P, Spizzichino A (1963) The Scattering of Electromagnetic Waves from Rough Surfaces. Pergamon (Republished by Artech, 1987), p 503.

Belmonte Rivas M, Martin-Neira M (2006) Coherent GPS Reflections From the Sea Surface. IEEE Geosci Remote Sens Lett 3:28–31. doi: 10.1109/LGRS.2005.855617

Bilich A, Axelrad P, Larson KM (2007) Scientific utility of the signal-to-noise ratio (SNR) reported by geodetic GPS receivers. Proc ION GNSS. Fort Worth, TX, pp 1999–2010.

Bilich A, Larson KM, Axelrad P (2008) Modeling GPS phase multipath with SNR: Case study from the Salar de Uyuni, Boliva. J Geophys Res 113:B04401. doi: 10.1029/2007JB005194

Bourlier C, Pinel N, Fabbro V (2006) Illuminated height PDF of a random rough surface and its impact on the forward propagation above oceans at grazing angles. 2006 First European Conference on Antennas and Propagation. IEEE, pp 1–6. doi: 10.1109/EUCAP.2006.4584894

Byun SH, Hajj GA, Young LE (2002) Development and application of GPS signal multipath simulator. Radio Sci 37:1098. doi: 10.1029/2001RS002549

Cardellach E, Ao CO, De la Torre Juárez M, Hajj GA (2004) Carrier phase delay altimetry with GPS-reflection/occultation interferometry from low Earth orbiters. Geophys Res Lett 31:4–7. doi: 10.1029/2004GL019775

Cardellach E, Fabra F, Rius A, Pettinato S, D'Addio S (2012) Characterization of dry-snow sub-structure using GNSS reflected signals. Remote Sens Environ 124:122–134. doi: 10.1016/j.rse.2012.05.012

Chen A, Chabory A, Escher A, Macabiau C (2009) Development of a GPS Deterministic Multipath Simulator for an Efficient Computation of the Positioning Errors. Proc ION GNSS. Savannah, GA, pp 2378–2390.

Chen A, Chabory A, Escher A, Macabiau C (2010) Hybrid Deterministic-Statistical GPS Multipath Simulator for Airport Navigation. In: Bonefačić D, Bosiljevac M (eds) IECom - 20th International Conference on Applied Electromagnetics and Communications. Dubrovnik, Croatia, pp 3–6.

Chen C-C, Gao S, Maqsood M (2012a) Antennas for Global Navigation Satellite System Receivers. In: Imbriale WA, Gao S, Boccia L (eds) Space Antenna Handbook. Wiley, pp 548–595. doi: 10.1002/9781119945147.ch14

Chen Q, Backen S, Akos DM, Larson KM (2013) GPS snow depth sensing with a monopole antenna. (in preparation)

Chen X, Parini CG, Collins B, Yao Y, Ur Rehman M (2012b) Antennas for Global Navigation Satellite Systems. Wiley, p 232. doi: 10.1002/9781119969518

Choi K, Bilich A, Larson KM, Axelrad P (2004) Modified sidereal filtering: Implications for high-rate GPS positioning. Geophys Res Lett 31:L22608. doi: 10.1029/2004GL021621

Cloude S (2009) Polarisation: Applications in Remote Sensing. Oxford University Press, New York, p 352.

Eamer J (2007) Global Outlook for Ice and Snow. United Nations Environment Programme, p 235 p.

Ercek R, De Doncker P, Grenez F (2005) Study of pseudo-range error due to non-line-of-sight-multipath in urban canyons. Proc ION GNSS. Long Beach, CA, pp 1083–1094.

Evans J, Capon J, Shnidman D (1989) Multipath Modeling for Simulating the Performance of the Microwave Landing System. Linc Lab J 2:459–474.

Fabra F, Cardellach E, Rius A, Ribo S, Oliveras S, Nogues-Correig O, Belmonte Rivas M, Semmling M, D'Addio S (2012) Phase Altimetry With Dual Polarization GNSS-R Over Sea Ice. IEEE Trans Geosci Remote Sens 50:2112–2121. doi: 10.1109/TGRS.2011.2172797

Falletti E, Pini M, Presti L Lo (2011) Low Complexity Carrier-to-Noise Ratio Estimators for GNSS Digital Receivers. IEEE Trans Aerosp Electron Syst 47:420–437. doi: 10.1109/TAES.2011.5705684

Fan K., Ding XL (2006) Estimation of GPS Carrier Phase Multipath Signals Based on Site Environment. J Glob Position Syst 5:22–28. doi: 10.5081/jgps.5.1.22

Fontana RD, Cheung W, Novak PM, Stansell TA (2001) The new L2 civil signal. Proc ION GPS. Institute of Navigation, Salt Lake City, UT, pp 617–631.

Georgiadou Y, Kleusberg A (1988) On carrier signal multipath effects in relative GPS positioning. Manuscr Geodaet 12:172–179.

Geren W, Murphy T, Pankaskie T (2008) Analysis of airborne GPS multipath effects using high-fidelity EM models. IEEE Trans Aerosp Electron Syst 44:711–723. doi: 10.1109/TAES.2008.4560216

Glenson S, Lowe S, Zavorotny V (2009) Remote sensing using bistatic GNSS reflections. In: Gleason S, Gebre-Egziabher D (eds) GNSS Applications and Methods. Artech, p 508 pp.

Gomez S, Panneton R, Saunders P, Hwu S, Lu B (1995) GPS multipath modeling and verification using geometrical theory of diffraction. Proc ION GPS. Palm Springs, CA, pp 195–204.

GPSD-USAF (2011) Navstar GPS Space Segment/Navigation User Segment Interfaces, Revision F (IS-GPS-200F). 210.

Gray DM, Male DH (1981) Handbook of Snow – Principles, Processes, Management and Use. Pergamon, p 776.

Groves PD, Jiang Z, Skelton B, Cross PA, Lau L, Adane Y, Kale I (2010) Novel multipath mitigation methods using a dual-polarization antenna. Proc ION GNSS. Portland, OR, pp 140–151.

Hannah BM, Walker RA, Kubik K (1998) Towards a complete virtual multipath analysis tool. Proc ION GPS. Nashville, TN, pp 1055–1063.

Hardy JR, Davis Y, Koh D, Cline K, Elder R, Armstrong HP, Marshall T, Painter GC, Saint-Martin R, DeRoo K, Sarabandi T, Graf T, Koike, McDonald K (2008) NASA Cold Land Processes Experiment (CLPX 2002/03): Local Scale Observation Site. J Hydrometeor 9:1434–1442.

Helm A, Montenbruck O, Ashjaee J, Yudanov S, Beyerle G, Stosius R, Rothacher M (2007) GORS – A GNSS Occultation, Reflectometry and Scatterometry Space Receiver. Proc ION GNSS. Fort Worth, TX, pp 2011–2021.

Irsigler M, Avila-Rodriguez JA, Hein GW (2005) Criteria for GNSS multipath performance assessment. Proc ION GNSS. Long Beach, CA, pp 2166–2177.

Jacobson MD (2010) Snow-covered lake ice in GPS multipath reception – Theory and measurement. Adv Space Res 46:221–227. doi: 10.1016/j.asr.2009.10.013

Kavak A, Vogel WJ, Xu G (1998) Using GPS to measure ground complex permittivity. Electron Lett 34:254–255. doi: 10.1049/el:19980180

Kelly REJ, Chang ATC, Tsang L, Foster JL (2003) A Prototype AMSR-E Global Snow Area and Snow Depth Algorithm. IEEE Trans Geosci Remote Sens 41:230–242.

King M a., Watson CS (2010) Long GPS coordinate time series: Multipath and geometry effects. J Geophys Res 115:B04403. doi: 10.1029/2009JB006543

Langley RB (1997) GPS receiver system noise. GPS World 8:40–45.

Larson KM, Gutmann ED, Zavorotny VU, Braun JJ, Williams MW, Nievinski FG (2009) Can we measure snow depth with GPS receivers? Geophys Res Lett 36:L17502. doi: 10.1029/2009GL039430

Larson KM, Nievinski FG (2013) GPS snow sensing: results from the EarthScope Plate Boundary Observatory. GPS Solut 17:41–52. doi: 10.1007/s10291-012-0259-7

Larson KM, Ray RD, Nievinski FG, Freymueller JT (2013) The Accidental Tide Gauge: A GPS Reflection Case Study From Kachemak Bay, Alaska. IEEE Geosci Remote Sens Lett 1–5. doi: 10.1109/LGRS.2012.2236075

Larson KM, Small EE, Gutmann ED, Bilich AL, Braun JJ, Zavorotny VU (2008) Use of GPS receivers as a soil moisture network for water cycle studies. Geophys Res Lett 35:L24405. doi: 10.1029/2008GL036013

Lau L, Cross P (2007) Development and testing of a new ray-tracing approach to GNSS carrier-phase multipath modelling. J Geod 81:713–732. doi: 10.1007/s00190-007-0139-z

Lemke PJ, Ren RB, Alley I, Allison J, Carrasco G, Flato Y, Fujii G, Kaser P, Mote RH, Zhang T, Thomas, T. (2007) Observations: Changes in Snow, Ice and Frozen Ground. In: Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt T and HLM (ed) Climate Change 2007: The Physical Science Basis. University Press, Cambridge, Mass., p 996 pp.

Lestarquit L, Nouvel O (2012) Determining and measuring the true impact of C/A code cross-correlation on tracking—Application to SBAS georanging. Proc IEEE/ION PLANS. IEEE, pp 1134–1140. doi: 10.1109/PLANS.2012.6236968

Lippincott W, Milligan T, Igli D (1996) Method for calculating multipath environment and impact on GPS receiver solution accuracy. Proc ION NTM. Santa Monica, CA, pp 707–722.

Löfgren JS, Haas R, Scherneck H-G, Bos MS (2011) Three months of local sea level derived from reflected GNSS signals. Radio Sci 46:1–12. doi: 10.1029/2011RS004693

López-Moreno JI, Latron J (2008) Influence of canopy density on snow distribution in a temperate mountain range. Hydrological Processes 22:117–126. doi: 10.1002/hyp.6572

Lowe S, Chaubell J, Hajj G (2007) A GNSS-reflections simulator and its application to widelane observations. 2007 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp 5101–5104. doi: 10.1109/IGARSS.2007.4424009

Macabiau C, Roturier B, Chatre E, Renard A (1999) Airport multipath simulation for siting DGPS reference stations. Proc ION NTM. San Diego, CA, pp 135–144.

Martin-Neira M, Colmenarejo P, Ruffini G, Serra C (2002) Altimetry precision of 1 cm over a pond using the wide-lane carrier phase of GPS reflected signals. Can J Remote Sens 28:394–403. doi: 10.5589/m02-039

Milligan TA (2005) Modern Antenna Design. Wiley-IEEE, p 632.

Misra P, Enge P (2006) Global positioning system: signals, measurements, and performance, 2nd ed. 2nd ed. Ganga-Jamuna, p 569.

Mora-Castro EJ, Carrascosa-Sanz C, Ortega G (1998) Characterisation of the Multipath Effects on the GPS Pseudorange and Carrier Phase Measurements. Proc ION GPS. Nashville, TN, pp 1065–1074.

NRC (2007) Earth science and applications from space: national imperatives for the next decade and beyond. National Research Council, Committee on Earth Science and Applications from Space, NAP

Ozeki M, Heki K (2011) GPS snow depth meter with geometry-free linear combinations of carrier phases. J Geod 86:209–219. doi: 10.1007/s00190-011-0511-x

Pany T (2010) Navigation Signal Processing for GNSS Software Receivers. Artech, p 352.

Park H, Marchan-Hernandez JF, Rodriguez-Alvarez N, Valencia E, Ramos-Perez I, Bosch-Lluis X, Camps A (2010) End-to-end simulator for Global Navigation Satellite System Reflectometry space mission. IEEE IGARSS. Honolulu, HI, pp 4294–4297. doi: 10.1109/IGARSS.2010.5650564

Ray JK, Cannon ME (2001) Synergy Between Global Positioning System Code, Carrier, and Signal-to-Noise Ratio Multipath Errors. J Guid Control Dynam 24:54–63. doi: 10.2514/2.4675

Rigden GJ, Elliott JR (2006) 3dM – A GPS Receiver Antenna Site Evaluation Tool. Proc ION NTM. Monterey, CA, pp 554–563.

Rignot E, Echelmeyer K, Krabill W (2001) Penetration depth of interferometric synthetic-aperture radar signals in snow and ice. Geophys Res Lett 28:3501–3504. doi: 10.1029/2000GL012484

Rodgers CE (1992) Multipath Simulation Software Developed for the Design of a Low Multipath DGPS Antenna for the US Coast Guard. Proc ION GPS. Albuquerque, NM, pp 43–50.

Rodriguez-Alvarez N, Bosch-Lluis X, Camps A, Aguasca A, Vall-llossera M, Valencia E, Ramos-Perez I, Park H (2011a) Review of crop growth and soil moisture monitoring from a ground-based instrument implementing the Interference Pattern GNSS-R Technique. Radio Sci 46:1–11. doi: 10.1029/2011RS004680

Rodriguez-Alvarez N, Camps A, Vall-llossera M, Bosch-Lluis X, Monerris A, Ramos-Perez I, Valencia E, Marchan-Hernandez JF, Martinez-Fernandez J, Baroncini-Turricchia G, Perez-Gutierrez C, Sanchez N (2011b) Land Geophysical Parameters Retrieval Using the Interference Pattern GNSS-R Technique. IEEE Trans Geosci Remote Sens 49:71–84. doi: 10.1109/TGRS.2010.2049023

Rott H, Yueh SH, Cline DW, Duguay C, Essery R, Haas C, Hélière F, Kern M, Macelloni G, Malnes E, Nagler T, Pulliainen J, Rebhan H, Thompson A (2010) Cold Regions Hydrology High-Resolution Observatory for Snow and Cold Land Processes. Proceedings of the IEEE 989:752–765.

Satirapod C, Rizos C (2005) Multipath mitigation by wavelet analysis for GPS base station applications. Survey Review 38:2–10. doi: 10.1179/003962605791521699

Semmling AM, Beyerle G, Stosius R, Dick G, Wickert J, Fabra F, Cardellach E, Ribó S, Rius A, Helm A, Yudanov SB, D'Addio S (2011) Detection of Arctic Ocean tides using interferometric GNSS-R signals. Geophys Res Lett 38:38–41. doi: 10.1029/2010GL046005

Small EE, Larson KM, Braun JJ (2010) Sensing vegetation growth with reflected GPS signals. Geophys Res Lett 37:L12401. doi: 10.1029/2010GL042951

Spilker JJ, Enge PK, Parkinson BW, Axelrad P (1996) Global Positioning System: Theory and Applications. AIAA, p 781.

Spirent (2010) Testing Multipath Performance of GNSS Receivers (MCD00136AAA). 22.

Treuhaft RN, Lowe ST, Zuffada C, Chao Y (2001) 2-cm GPS altimetry over Crater Lake. Geophys Res Lett 28:4343. doi: 10.1029/2001GL013815

Ulaby FT, Moore RK, Fung AK (1986) Microwave Remote Sensing: Active and Passive: Radar Remote Sensing and Surface Scattering and Emission Theory. Artech House, p 608.

Veatch W, Brooks PD, Gustafson JR, Molotch NP (2009) Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site. Ecohydrology 2:115–128. doi: 10.1002/eco.45

Weiss J, Axelrad P, Anderson S (2007) A GNSS Code Multipath Model for Semi-Urban, Aircraft, and Ship Environments. Navig 54:293–307.

WMO (2007) A cryosphere theme report for the IGOS partnership,. WMO/TD

Woo KT (2000) Optimum Semicodeless Carrier-Phase Tracking of L2. Navig 47:82–99.

Zavorotny VU, Larson KM, Braun JJ, Small EE, Gutmann ED, Bilich AL (2010) A Physical Model for GPS Multipath Caused by Land Reflections: Toward Bare Soil Moisture Retrievals. IEEE J Sel Topics Appl Earth Observ 3:100–110. doi: 10.1109/JSTARS.2009.2033608

Zavorotny VU, Voronovich AG (2000) Scattering of GPS signals from the ocean with wind remote sensing application. IEEE Trans Geosci Remote Sens 38:951–964. doi: 10.1109/36.841977

Zhu Z, van Graas F (2009) Earth-Surface Multipath Detection and Error Modeling for Aircraft GPS Receivers. Navig 56:45–56.